COMPARISONS OF APPROACHES TO MODELLING TREE TAPER, STAND STRUCTURE AND STAND DYNAMICS IN FOREST PLANTATIONS

A thesis
submitted in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy in the
University of Canterbury

by

Ricardo Methol

University of Canterbury
2001

New Zealand School of Forestry

Abstract

This study addressed two general tree modelling topics: (i) taper, and (ii) growth and yield. Several modelling options were considered and evaluated in detail for each of the two topics. The use of various datasets (two and three, respectively) enabled sound conclusions to be drawn on a variety of methodological aspects.

In the first part of the study, several existing taper equations of different types were compared, and new variants were proposed. Variable-exponent equations exhibited the maximum levels of accuracy in terms of estimating diameter, height and volume. A variant of the classic segmented taper equation (Max and Burkhart 1976) was found to be best. This equation showed reasonably accurate diameter and volume estimations and was able to be inverted for height estimations and explicitly integrated for volume estimations. Bark models that predict under-bark diameters from over-bark diameters, *dbh* and total tree height were developed and used in a 2-step composite approach for predicting under-bark diameters. This approach has been proposed to minimise the number of under-bark measurements when collecting taper data (Gordon *et al.* 1995), but its validity has not been proved. The results obtained in this study showed that this composite approach was reliable for the two datasets used in this study. Practical suggestions for collecting taper data when adopting this approach were provided.

In the second part, three approaches for developing growth and yield models with comparable output resolution were compared. These approaches included: (i) diameter distribution models; (ii) relative-basal-area-based dis-aggregative approaches; and (iii) individual tree models. To facilitate comparisons, tree-level outputs from the latter two approaches were grouped by diameter class. Within each approach, a variety of equations and procedures was evaluated in order to obtain the best possible model for each approach. Diameter distribution models based on the reverse Weibull distribution with a parameter-recovery approach provided acceptable depictions of actual stand tables, and

were more accurate than other methods over long projections. Nonetheless, the best representation of actual diameter distributions was achieved with individual-tree models. Between those modelling approaches that provided compatible tree-level and stand-level output resolution, the adjusted individual-tree model (ITM_{adj}) approach exhibited lower error indices than the relative-basal-area-based dis-aggregative approach. This was attributed to the increased complexity of diameter increment equations, which relied on stand-level and tree-level variables, as opposed to the simple formulation of relative basal area equations.

Adjusted individual-tree models were preferred to unadjusted individual-tree models (ITM) for all three species studied. For the New Zealand grown species (*Pinus radiata* and *Pseudotsuga menziesii*), this decision was based on (i) more accurate basal area and stocking estimations from stand-level models as compared to basal area and stocking estimations derived by aggregating tree-level projections; and (ii) the lower error indices for depicting diameter distributions that were achieved with the ITM_{adj} as compared to the ITM. For *E. grandis* grown in Uruguay, however, the main criterion for choosing the ITM_{adj} was compatibility with the stand-level model, given that the ITM produced slightly lower error indices than the ITM_{adj} for depicting diameter distributions.

Complete growth and yield models were built for all three species for the Central North Island of New Zealand and for Zones 7, 8 and 9 of Uruguay. These provided a range of options for performing forecasts of future stand conditions, depending on the type of input data available. Model options include fully compatible whole-stand, diameter-distribution and individual-tree models. User-friendly computer simulators including all these options were developed. These simulators allow comparisons of thinning schedules and rotation ages, providing information on the main stand variables, as well as tree-level (or diameter class) details and merchantable volumes by log type.

Regional models for estimating individual-tree heights from diameter and mean top height plot measurements, which form part of the growth models just described, can also be used independently. Their use would allow resources to be saved by minimising the amount of inventory height measurements.

Contents

PART II - MODELLING STAND STRUCTURE AND DYNAMICS

List of Tables

List of Figures

Preface

This thesis is divided into three parts. Part I (Chapters 1 to 5) documents the assessment of numerous taper models. Models for predicting under-bark diameters from over-bark diameters (bark models) are also described, along with an evaluation of the use of these models in a two-step composite modelling approach.

Part II (Chapters 6 to 12) constitutes the core of this thesis, describing the development, evaluation and comparison of modelling approaches with various levels of resolution (stand, diameter class and tree). This part is largely independent to the previous part, although a tree volume equation derived from a taper model developed in Part I was used for calculating tree and stand volumes.

Part III (Chapter 13 and 14) discusses and summarises the key elements in the preceding sections. Chapter 13 contains a general discussion of all topics covered in the study, while Chapter 14 briefly reports the main conclusions of this thesis.

PART I

TAPER AND BARK MODELLING

CHAPTER 1

INTRODUCTION

Good forest management requires accurate information on the current growing stock and future growth potential. The former is obtained through forest inventories and the latter is estimated or projected from a current inventory by growth and yield models. An accurate forest inventory requires per hectare volume estimations (either total or merchantable) calculated from the summation of individual tree volumes. Tree volume equations can be used for this task, although the recent trend is to use more flexible *taper estimation systems*. Accurate taper functions are essential for pre-harvest inventory, when stems are assessed and then summarised by log type.

Taper estimation systems are based on a taper function that describes the tree profile and can provide:

(i) diameter estimates (either under- or over bark) at any point of the stem;
(ii) estimates of heights at which a given diameter occurs along the stem;
(iii) estimates of total stem volumes (either under- or over bark);
(iv) estimates of merchantable volumes (either under- or over bark) to any merchantable height or minimum upper-stem diameter and from any stump height; and
(v) estimates of individual log volumes.

A major part of the research concerning taper modelling since studies by Daemerschalk (1972) has addressed the compatibility aspect of taper and tree volume functions. Compatibility had major relevance in the past, when explicit tree volume equations were required to make volume estimations in a simple manner. Nowadays, the development of fast computers has made explicit tree volume functions less relevant. Volume calculations of any sort (e.g. total volume, merchantable volume, log volume assortments) can be performed by integration of an appropriate taper function. Some taper functions can be integrated yielding an explicit expression of tree volume as a function of tree height and DBH. Others cannot, but numerical integration is still available (e.g. by calculating diameters and lengths and then estimating volumes by Smalian's, Newton's or conoid formulae).

Taper equations developed in most taper studies have been concerned with the under-bark stem profile (Bruce *et al.* 1968; Demaerschalk 1972, 1973; Ormerod 1973; Goulding and Murray 1976; Max and Burkhart 1976; Amidon 1984; McClure and Czaplewski 1986; Hayward 1987; Alemdag 1988; Kozak 1988; Candy 1989; Perez *et al.* 1990; Allen *et al.* 1992; Newnham 1992; Gál and Bella 1995; Kozak 1997). Some researchers have modelled both under- and over-bark tree taper (Cao *et al.* 1980; Byrne and Reed 1986; Bailey 1994; Figueiredo-Filho *et al.* 1996; Gordon *et al.* 1999) and only a few researchers have modelled only over-bark taper (Reed and Byrne 1985). Having information on over-bark diameters in addition to under-bark diameters and volumes might not be crucial when the only management objective deals with typical timber forest products. However, information on the amount of bark can be relevant for total biomass estimations and carbon sequestration accounting, which are becoming the driving forces in many recent large-scale afforestation programmes.

1.1 CLASSIFICATION OF TAPER MODELS

Taper functions reported in the literature can be grouped into four categories:

a) single functions, including polynomials, that represent the whole bole (Bruce *et al.* 1968; Ormerod 1973; Hilt 1980; Gordon *et al.* 1995)

b) segmented polynomials that are smoothly joined and each segment represents different parts of the bole (Max and Burkhart, 1976; Cao *et al.* 1980; Temu, 1992)

c) within-tree variable form (or variable exponent) equations (Kozak, 1988; Newnham, 1992)

d) between-tree variable form functions (Reed and Byrne 1985; Newberry and Burkhart 1986; Real and Moore 1988; Candy 1989; Allen *et al.* 1992)

While the first two types are better known the last two have been relatively less investigated. The best-known taper models of each category are described in the following sections.

Unless otherwise stated, the terms X, *y* and *z* represent:

$$X = \frac{H-h}{H-BH} \; ; \quad y = \frac{H-h}{H} ; \text{ and } z = \frac{h}{H}$$

where BH is breast height (1.30 m in Uruguay and 1.40 m in New Zealand); H is total tree height (m) and h is the length (m) from ground to upper-stem diameter d (cm) ($0 \leq h \leq H$).

1.1.1 Single functions

1.1.1.1 Simple taper models with 2, 3 or 4 parameters

Many taper models were derived from tree volume equations so as to be *compatible* with them. The theory behind the derivation of a taper model from a tree volume equation, first reported by Demaerschalk (1972), was remarkable at that time. Before Demaerschalk's study it was common that tree volume and taper equations were both in use for a given population, and that volumes obtained from the tree volume equation were not equal to the volumes obtained by integration of the taper equation. Computers were not widespread in those days, and simple, explicit volume equations were

necessary. Equation 1.1 was derived by Demaerschalk (1973) from the simple form-factor tree volume equation ($V=f.g.H$, where $g=dbh^2\pi/4$), whereas Equation 1.2 was derived by Reed and Green (1984) from the Schumacher's tree volume equation ($V=a.dbh^bH^c$).

$$d = \sqrt{b_1 dbh^2 \left(\frac{H-h}{H}\right)^{b_2}} \tag{1.1}$$

$$d = b_1 dbh^{b_2} \frac{(H-h)^{b_3}}{H^{b_4}} \tag{1.2}$$

Other taper equations were derived from classic tree volume equations such as the *logarithmic* (Demaerschalk, 1972), volume over basal area as a function of height and squared height ($V/g=a+bH+cH^2$), form factor with intercept (Demaerschalk, 1973), Honer's (1965) and the combined variable (Demaerschalk, 1973; Reed and Green, 1984). For the purpose of this study it was deemed sufficient to select just two models of this type for the list of candidate models. In preliminary analyses Equations 1.1 and 1.2 exhibited greater levels of accuracy and precision for diameter estimations than other models and converged easily when fitted by non-linear regression.

Tree volume equations based on *dbh* and height with only two or three parameters can be very accurate, although tree taper is more difficult to model. The simplicity of taper equations derived from tree volume equations makes them less flexible and powerful than taper models specifically conceived to describe stem profiles.

Ormerod (1973) presented a simple model based on geometric principles (Equation 1.3). The equation is so conditioned that when h=H, d=0. As the model was conceived to describe under-bark taper, predicted diameters when h=BH are given by a constant proportion of *dbh* (i.e. b_1). By adding the term $b_0(H-h)$ that vanishes at tree top and compensates for the double-bark thickness at BH proportionally to tree size (i.e. H) it is possible to utilise this model to describe over-bark taper as well, without losing the condition of zero diameter at tree top (Equation 1.3a).

$$d = b_1 dbh\left(\frac{H-h}{H-BH}\right)^{b_2} \quad (1.3)$$

$$d = b_0(H-h) + b_1 dbh\left(\frac{H-h}{H-BH}\right)^{b_2} \quad (1.3a)$$

The parameter b_1 in Equation 1.3 has been referred to as the *taper* parameter (Prodan *et al.* 1997). However, as Newberry and Burkhart (1986) pointed out, $b_1 dbh$ is the real taper *parameter,* accounting for tree taper not accounted for by tree *dbh*. The parameter b_2 in equation 1.3 is the *form* parameter. The terms taper and form can be regarded as synonymous when referring to the stem profile. Newberry and Burkhart (1986) quoted the following definitions for taper and form given by Gray[1]: (i) form is the characteristic shape of a solid, and (ii) taper is the rate of narrowing in diameter in relationship to the increase in height.

Amidon (1984) proposed a simple and *stable* model to represent tree profiles (Equation 1.4). It has the desirable property of yielding zero diameter at tree top. This model also predicts *d* at breast height as a fixed proportion of *dbh*. Although this property can be reasonable when predicting under-bark diameters, it prevents using the model for over-bark diameters (when predicted diameter at breast height should equal *dbh*). As in the case of Ormerod's model, the addition of the term $b_0(H-h)$ would allow using this model to describe both under- and over-bark taper (Equation 1.4a).

$$d = b_1 dbhX + b_2 \frac{(H^2 - h^2)(h - bh)}{H^2} \quad (1.4)$$

$$d = b_0(H-h) + b_1 dbhX + b_2 \frac{(H^2 - h^2)(h - bh)}{H^2} \quad (1.4a)$$

Trigonometric taper equations have been proposed by Thomas and Parresol (1991), who pointed out the direct analogy between the relative-height/relative-diameter plots and trigonometric functions on the unit circle. After carefully examination of candidate

[1] Gray, H.R. 1956. The form and taper of forest-tree stems. Oxford Univ., Imp. For. Inst. Pap. No. 32. pp. 1-74

trigonometric functions, Thomas and Parresol (1991) proposed the model here referred to as Equation 1.5 and described it as a simple, flexible, trigonometric taper equation.

$$d = \sqrt{dbh^2\left[b_1(z-1) + b_2\sin(c.z.\pi) + \frac{b_3}{\tan(\pi z/2)}\right]} \qquad (1.5)$$

1.1.1.2 Polynomials on relative height

Polynomials on some measure of relative height (i.e. X, y, or z) have been proposed to describe the typical curve that appears when the ratios of d/*dbh* or d^2/dbh^2 are plotted against relative height. These curves are exemplified below (Figure 1.1) from a random sample of 20% of the observations in the *E. grandis* data used in this study.

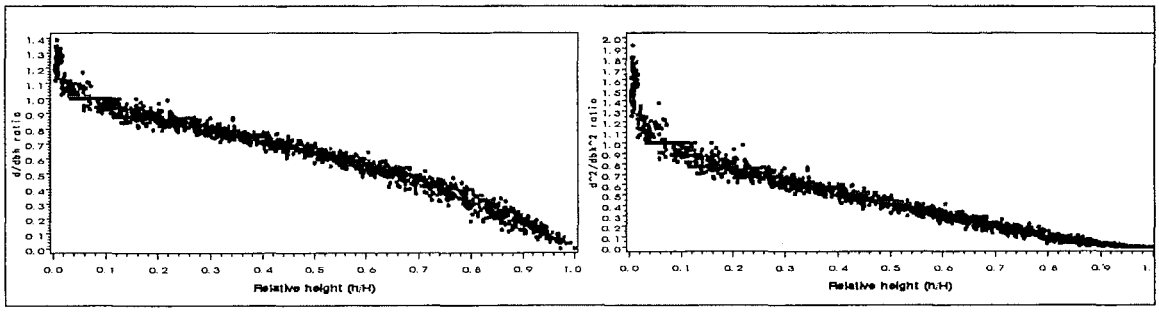


Figure 1.1 Diameter over bark/dbh versus relative height (left); and sectional area relative to sectional area at BH (d^2/dbh^2) versus relative height (right).

A simple polynomial on relative height (z) restricted to three terms (Equation 1.6) was proposed by Kozak *et al.* (1969).

$$d = dbh\sqrt{a + b\frac{h}{H} + c\left(\frac{h}{H}\right)^2} \qquad (1.6)$$

Instead of describing the ratio of sectional area to sectional area at BH (as Equation 1.6) the so-called fifth-degree polynomial (Equation 1.7) depicts the ratio of diameter to *dbh* (Figueired-Filho *et al.* 1996).

$$d = dbh\left(b_0 + b_1 z + b_2 z^2 + b_3 z^3 + b_4 z^4 + b_5 z^5\right) \qquad (1.7)$$

A more complex polynomial taper equation was developed by Bruce *et al.* (1968) (Equation 1.8). Instead of h/H (i.e. z) they used X as the descriptor of relative height, which vanishes at tree top and equals one at breast height. As in the case of Equations 1.3 and 1.4, Equation 1.8 predicts under-bark diameter at breast height as a fixed proportion of *dbh*. This proportion is represented by the value of the coefficient b_1. Again, this condition prevents using the model for over-bark diameters and it can also be overcome by adding the term $b_0(H-h)$ (Equation 1.8a). Another variant in Equation 1.8a is that a parameter *c* to be estimated from the data substituted the exponent 3/2.

$$\frac{d^2}{dbh^2} = b_1 X^{\frac{3}{2}} + b_2\left(X^{\frac{3}{2}} - X^3\right)dbh + b_3\left(X^{\frac{3}{2}} - X^3\right)H + \tag{1.8}$$
$$b_4\left(X^{\frac{3}{2}} - X^{32}\right)H.dbh + b_5\left(X^{\frac{3}{2}} - X^{32}\right)H^{\frac{1}{2}} + b_6\left(X^{\frac{3}{2}} - X^{40}\right)H^2$$

$$\frac{d^2}{dbh^2} = b_0(H-h) + b_1 X^c + b_2\left(X^c - X^3\right)dbh + b_3\left(X^c - X^3\right)H + \tag{1.8a}$$
$$b_4\left(X^c - X^{32}\right)H.dbh + b_5\left(X^c - X^{32}\right)H^{\frac{1}{2}} + b_6\left(X^c - X^{40}\right)H^2$$

A variation of Equation 1.8 without the parameter b_1 was later proposed by Hilt (1980) and is presented here as Equation 1.9. Being developed to describe under-bark taper, the diameter under bark at breast height (*dbhib*) needed to be used instead of *dbh* (because of the absence of the parameter b_1). However, if the model is to be fitted to over-bark data then *dbh* could be used. In order to utilise this model to describe both under and over-bark taper, with *dbh* (instead of *dbhib*) and H as independent variables, the term $b_0(H-h)$ and the parameter b_1 for the term in $X^{3/2}$ should be added (Equation 1.9a). As in Model 8a, a parameter *c* to be estimated from the data substituted the exponent 3/2 in Model 1.9a.

$$\frac{d^2}{dbhib^2} = X^{\frac{3}{2}} + b_2\left(X^{\frac{3}{2}} - X^3\right)H + b_3\left(X^{\frac{3}{2}} - X^3\right)dbh.H \tag{1.9}$$
$$+ b_4\left(X^{\frac{3}{2}} - X^{30}\right)dbh + b_5\left(X^{\frac{3}{2}} - X^{30}\right)dbh.H$$

$$\frac{d^2}{dbh^2} = b_0(H - h) + b_1X^c + b_2(X^c - X^3)H + b_3(X^c - X^3)dbh.H \quad (1.9a)$$
$$+ b_4(X^c - X^{30})dbh + b_5(X^c - X^{30})dbh.H$$

A model similar to Equations 1.8 and 1.9 was used to model tree taper on a tree-by-tree basis (Real and Moore 1988). In the mentioned study, the three parameters of Equation 1.10 were regressed on tree variables such as H, crown ratio and shape quotient [H/(*dbh*/12)]. Although this model can be very well suited to the between-tree variable form approach (see Section 1.1.1.4), its simple formulation makes it less versatile than Equations 1.8 and 1.9 when fitted to a group of trees.

$$d = \sqrt{dbh^2\{X^2 + b_1(X^3 - X^2) + b_2(X^8 - X^2) + b_3(X^{40} - X^2)\}} \quad (1.10)$$

In order to make Equation 1.10 suited for both under- and over-bark applications the following variant could be appropriate (Equation 1.10a):

$$d = \sqrt{dbh^2\{b_0X^2 + b_1(X^3 - X^2) + b_2(X^8 - X^2) + b_3(X^{40} - X^2) + b_4(H - h)\}} \quad (1.10a)$$

1.1.2 Segmented taper models

Different segments of the stem are commonly assumed to approximate various geometric solids. The lower bole portion is generally assumed to be a neiloid frustum, the middle portion a paraboloid frustum, and the upper portion a cone (e.g. Husch *et al.* 1972). This led Max and Burkhart (1976) to model each segment individually and join or spline these segments at points defined by join-point parameters to form a single model. In this way a different kind of taper model was created, the *segmented* taper model.

Besides being the first application of segmented polynomial theory to taper modelling, the model proposed by Max and Burkhart (1976) is probably the most popular taper equation. It is frequently used as a benchmark with which to compare new taper models (Cao *et al.* 1980; Amidon 1984; Byrne and Reed 1986; Real and Moore 1988; Thomas

and Parresol 1991; Newnham 1992; Kozak and Smith 1993; Figueiredo-Filho *et al.* 1996; Muhairwe 1999). Although the dependent variable in the original formulation is the ratio of sectional area to sectional area at BH (i.e. d^2/dbh^2) the model can be arranged to predict *d* directly (Equation 1.11).

$$d = \sqrt{dbh^2\left[b_1\left(\frac{h}{H}-1\right)+b_2\left(\frac{h^2}{H^2}-1\right)+b_3\left(a_1-\frac{h}{H}\right)^2 I_1+b_4\left(a_2-\frac{h}{H}\right)^2 I_2\right]} \quad (1.11)$$

where $I_i=1$ if $z<a_i$ and $I_i=0$ otherwise

Cao *et al.* (1980) proposed a *compatible* segmented equation that also depends on total tree volume (V). As it had been decided to develop pure taper functions without any constraint in the present study, this model as such was not a candidate. However, if we consider that $V=c.dbh^2H$ (the form factor tree volume equation), then a pure taper equation (Equation 1.12) can be derived as in Byrne and Reed (1986).

$$d = \sqrt{dbh^2 c\left[2y+b_1\left(3y^2-2y\right)+b_2(y-a_1)^2 I_1+b_3(y-a_2)^2 I_2\right]} \quad (1.12)$$

$I_i=1$ if $y>=a_i$ and $I_i=0$ otherwise

c is a re-parametization of c/k, where k=0.00007854

Another segmented taper model, also derived from Cao's compatible taper equation, was proposed by Temu (1992). It has one parameter more than Equations 1.11 and 1.12 and was the preferred model in a taper study for Douglas-fir in New Zealand (Equation 1.13).

$$d = \sqrt{dbh^2 f\left[b_1 y^2+b_2 y+dbh.b_3(y-a_1)^2 I_1+dbh.b_4(y-a_2)^2 I_2\right]} \quad (1.13)$$

where $I_i=1$ if $y>=a_i$, $I_i=0$ otherwise

1.1.3 Variable-exponent taper equations

Within-tree variable form or variable-exponent equations are continuous functions describing the shape of the bole, with a changing exponent from ground to top to compensate for the neiloid, paraboloid and conic forms in different sections of the tree (Kozak 1988).

The first variable exponent model reported in the literature was proposed by Kozak (1988), and is here referred to as Equation 1.14.

$$d = a_0 dbh^{a_1} a_2^{dbh} \left(\frac{1-\sqrt{z}}{1-\sqrt{p}} \right)^{b_1 z^2 + b_2 \ln(z+0.001) + b_3 \sqrt{z} + b_4 \exp(z) + b_5 (dbh/H)} \quad (1.14)$$

where p is a parameter representing the lower inflection point of the stem profile curve.

Aiming at improving the accuracy of total tree volume estimations, Kozak (1997) developed a variant of Equation 1.14, which was referred to as the 1994 equation (Equation 1.15):

$$d = a_0 dbh^{a_1} a_2^{dbh} \left(\frac{1-\sqrt{z}}{1-\sqrt{0.01}} \right)^{b_1 z^{1/4} + b_2 z^{1/3} + b_3 z^{1/2} + b_4 \arcsin\left[1-\sqrt{z}\right] + b_5 (H/dbh+z) + b_6 H} \quad (1.15)$$

The variables in the exponent of Kozak's 1994 equation were very highly inter-correlated (multicollinearity), even more than those in the 1988 equation (Kozak, 1997). When severe multicollinearity exists the following problems occur (Myers 1990): (i) small changes in the data can produce significant changes in the parameter estimates (regression coefficients); (ii) regression coefficients have high standard errors, which affect the significance level of the corresponding independent variable; and (iii) the regression coefficients may have the wrong sign and (or) an unreasonable magnitude. The high level of multicollinearity of the 1994 equation led Kozak (1997) to develop a new equation (the 1995 equation) in which the variables of the exponent were selected so as to minimise the level of multicollinearity (Equation 1.16):

$$d = a_0 dbh^{a1} H^{a_2} X_K^{b_1 x^{0.1} + b_2 z^4 + b_3 \arcsin(1-\sqrt{z}) + b_4/\exp(dbh/H) + b_5 dbh^{X_K}} \quad (1.16)$$

where $X_K = \left(\frac{1-\sqrt{z}}{1-\sqrt{\frac{BH}{H}}} \right)$

Recently, a variant of Kozak's variable exponent model has been proposed by Muhairwe (1999). The main difference with Kozak's model is that the base of the exponential term is simply $1-\sqrt{z}$. In this way, the estimation of the inflexion point (p) is avoided. Muhairwe's model is presented as Equation 1.17.

$$d = a_0 dbh^{a1} a_2^{dbh} \left(1-\sqrt{z}\right)^C \quad (1.17)$$

where C is the exponent term containing various independent variables.

1.1.4 Between-tree variable form models

Between-tree variable form functions, also referred to as individual tree taper systems (Real and Moore 1988) are developed in two stages. In the first stage, a taper function in which parameters are easily interpretable (e.g. Ormerod 1973; Cao *et al.* 1980) is fitted to each tree. In the second stage, these parameters are regressed on tree and stand variables.

This approach was tried in preliminary investigations for the present study. However, its usefulness was deemed limited for a number of reasons. Firstly, taper models suited to this method need to be simple and their parameters *stable* and interpretable. By stable is meant that the parameters should always have the same sign and a value of the same level of magnitude. By interpretable is meant that the magnitude of each parameter can be somehow associated with some tree characteristics. This condition invalidated the use of complex formulations such as variable exponent models, which have proven to be very valuable. Secondly, fitting the models by non-linear regression on a tree-by-tree basis resulted in convergence failures for some trees. In general these trees had a reduced

number of observations and thus tended to be small trees. If this approach was to be followed and these trees disregarded for the second step, the developed models would had been biased for the whole population as a result of the under-representation of smaller trees. Thirdly, this 2-step approach is particularly suited for cases where more tree characteristics (e.g. crown ratio, mean annual diameter increment) and/or stand characteristics (stocking, mean top height, age) are available. That was not the case in the datasets available for the study. Furthermore, even when this information is available for research purposes, it is seldom recorded in standard inventory practice, and the models so developed would tend to be impractical.

1.2 EVALUATING AND COMPARING TAPER MODELS

Taper studies have proliferated in the last three decades (Fang and Bailey 1999) indicating the considerable importance of the topic. Because of the flood of taper equations, it is often difficult for managers to ascertain which is best suited for their purposes (Kozak and Smith 1993). This confusion may also apply to researchers fitting taper functions for new populations. Kozak and Smith (1993) discussed some standards to choose among several functions.

The first decision to be made when evaluating and comparing taper models involves the selection of criteria to be used. It is clear that the first screening of the candidate models need not be as detailed as the final selection of the best model from a reduced number of top candidates. For the first screening, overall statistics may suffice. On the other hand, more detailed analyses should be performed for the final selection of the best model. If the taper equation is to be used as a *taper estimating system*, its use will not only involve diameter predictions, but also heights at different merchantable diameters and merchantable volumes. The performance of the taper estimation system needs to be tested for all these variables. Furthermore, diameter projections should be evaluated at different positions along the stem (e.g. percent heights) and merchantable height and volume projections should be evaluated by *dbh* classes (Kozak and Smith, 1993).

Residual plotting is another useful way to ascertain the performance of models. It allows detection of trends of residuals in relation to predicted values, independent variables and other relevant variables not included in the model.

In order to determine the statistics most used to evaluate and compare taper models, a review of the criteria used in a number of published studies was carried out. All selected studies involved comparisons between taper models (Table 1.1).

Table 1.1 Statistics used for model comparisons in various published taper studies.

Reference	Bias	SEE	MAD	FI	SDR	RSS	RMSE	SSRR	MSE	Number
Demaerschalk (1972)		x								1
Demaerschalk (1973)		x								1
Ormerod (1973)		x								1
Newnham (1992)						x				1
Petersson (1997)	x									1
Max and Burkhart (1976)	x		x							2
Amidon (1984)	x	x								2
Thomas and Parresol (1991)		x		x						2
Allen *et al.* (1992)	x						x			2
Kozak and Smith (1993)	x	x								2
Bailey (1994)	x					x				2
Gal and Bella (1995)	x	x								2
Kozak (1997)	x	x								2
Fang and Bailey (1999)	x	x								2
Cao *et al.* (1980)	x		x		x					3
Newberry and Burkhart (1986)	x		x				x			3
Candy (1989)				x		x	x			3
Perez *et al.* (1990)				x		x			x	3
Byrne and Reed (1986)	x		x	x	x					4
Figueiredo-Filho *et al.* (1996)	x		x		x			x		4
Muhairwe (1999)	x	x	x	x				x		5
Total	14	10	6	5	3	4	3	2	1	
Percentage	67%	48%	29%	24%	14%	19%	14%	10%	5%	

Notes: Bias: mean residual; MAD: mean absolute difference; FI: fit index or percentage of the variance explained by the model; SDR: standard deviation of the residuals; RSS: residual sum of squares; RMSE: square root of the MSE (mean square error or residual mean square); SSRR: sum of the squared relative residuals.

Bias (mean residual) was the most frequently used statistic and is usually considered as a measure of *accuracy*. It is simply the average residual, that is, the average difference between actual (observed) and predicted values of the variable being modelled:

$$\text{Bias} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)}{n} \qquad (1.18)$$

The second most used statistic was the standard error of estimate (SEE), which is a measure of *precision* and complements bias very well. The SEE is usually[2] defined as follows:

$$SEE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-p}} \tag{1.19}$$

where p in the number of parameters.

The mean absolute deviation was the third most commonly used statistic, representing the average error of prediction (regardless of sign):

$$MAD = \frac{\sum_{i=1}^{n}\left|Y_i - \hat{Y}\right|}{n} \tag{1.20}$$

1.3 AUTOCORRELATED NATURE OF TAPER DATASETS

Data used for taper modelling have a major drawback from the statistical point of view: they are extremely autocorrelated (or serially correlated). This is because many measurements are usually taken from each tree. Fitting a model, either by linear or non-linear regressions, to autocorrelated data violates the statistical assumption that error terms are independent. This fact has no serious consequences on parameter estimation (parameter estimates are unbiased and consistent). However, autocorrelated data may have the following consequences pertaining to the statistical inferences that can be drawn (Kozak 1997):

- the estimators of the regression coefficients no longer have the minimum variance property

[2] Some authors define SEE with *n* as the denominator instead of *n-p* (e.g. Gal and Bella, 1995). This

- the calculated mean squared error (MSE) may underestimate the real variance of the residuals, while the standard errors of the regression coefficients may seriously underestimate the true standard deviations of the coefficients

- statistical tests using t or F distributions and confidence intervals are no longer reliable

Recently, statistical packages such as SAS have introduced new procedures capable of accounting for autocorrelated data. These procedures not only model the *fixed* effects (i.e. parameter estimation) but also the *random* effects (variance and covariance structures). From this perspective, models are referred to as *mixed* models. Developed originally for linear models, mixed models have since been expanded to incorporate generalised linear models and non-linear models (Candy 1997). Parameter estimates in PROC MIXED are obtained by restricted maximum likelihood (REML) methods implemented with a Newton-Raphson algorithm (SAS Institute Inc. 1996). Despite the availability of these techniques, there are very few published taper studies (e.g. Tassisa and Burkhart 1998) where the autocorrelated nature of the data has been effectively accounted for.

1.4 BARK MODELS

When modelling both under- and over-bark tree profiles, two general approaches can be followed. One is to model both independently and the other is to model the over-bark profile first and link the predicted diameter over bark *(dob)* to its corresponding diameter under bark *(dub)* through a *bark model*. The latter approach was proposed by Gordon *et al.* (1995) and was referred to as a *composite* approach. The purpose of this approach is to minimise the quantity of under-bark measurements (that are time consuming and often inaccurate) for developing taper functions. This approach was later followed in another taper study on *Eucalyptus* species by the same authors (Gordon *et al.* 1999), but its validity has not heretofore been studied.

Gordon (1983) developed the following non-linear model for predicting the bark thickness of *P. radiata* in New Zealand (Equation 1.21):

---

definition seems more suited to evaluate or validate models with an independent dataset.

$$\ln(B/dob) = b_0 + b_1\left(1-\frac{h}{H}\right)^{b_2} + b_3\left(\frac{h}{H}\right)^{b_4 H} + b_5 dbh + b_6\frac{H}{dbh} \quad (1.21)$$

where B is the double bark thickness in cm (i.e. dob-dub). This equation can be re-arranged, to predict *dub* directly (Equation 1.22):

$$dub = dob - \exp\left[b_0 + b_1\left(1-\frac{h}{H}\right)^{b_2} + b_3\left(\frac{h}{H}\right)^{b_4 H} + b_5 dbh + b_6\frac{H}{dbh}\right]dob \quad (1.22)$$

A simpler version of Equation 1.22 was later used by Penman (1988) in a study comparing the taper and bark thickness of radiata pine seedlings and cuttings. Bailey (1994) proposed a model to predict double bark thickness at any point of the tree, from total height and outside-bark values of *dbh*, stump diameter, and upper-stem diameter. As consistent measures of stump diameter were not available in the data available (nor are they likely to be recorded in standard inventory practice), this model was not evaluated. Gordon *et al.* (1995) developed the following bark model for *Eucalyptus saligna* in New Zealand:

$$dub = \sqrt{dob^2\left(a_0 + a_1 y + a_2 y^8\right)} \quad (1.23)$$

where y = (H-h) / H

The same model, but with the higher term restricted to 2, was fitted to the stringybark *Eucalyptus* species by Gordon *et al.* (1999). These polynomials describe the curve of the ratio of under-bark sectional area to over-bark sectional area that appears when plotted against *y*. This curve is exemplified below for both datasets available for the present study (Figure 1.2).

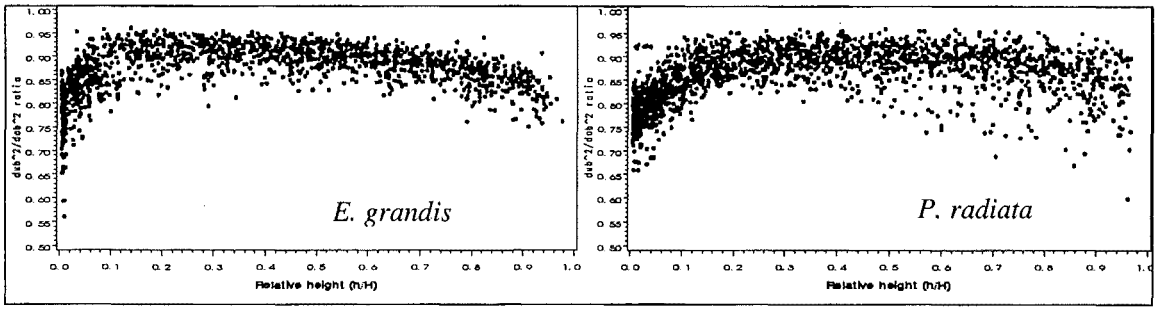


Figure 1.2 Ratio of under-bark to over-bark sectional area *versus* relative height (z) for *E. grandis* and *P. radiata*

1.5 STUDY OBJECTIVES

The objectives of this part of the study were:

(i) to evaluate and compare different types of taper models identifying their main advantages and drawbacks from practical and statistical points of view;

(ii) to evaluate the validity of the composite approach to develop over- and under-bark taper equations;

(iii) to select from among the range of published equations the best taper models for *E. grandis* in Uruguay and for *P. radiata* in the Auckland and Northland regions of New Zealand. In the case of *E. grandis*, the taper estimation system developed here was required for the growth models described in Part II of this study.

CHAPTER 2

DATA, METHODS AND PROCEDURES

2.1 DATA DESCRIPTION

The data available for this study are summarised in Table 2.1. Data for *E. grandis* were from three regions of Uruguay referred to as Zone 7, Zone 8 and Zone 9. Each region was represented by various stands (21, 8 and 18, respectively) covering a wide range of ages, site qualities and tree sizes. In each stand 8 trees on average were sampled. Over-bark diameters of *E. grandis* were measured at 0.15, 0.65, 1.30 m and at 1-m intervals thereafter up to the top of the tree. On average, 21 measures were taken from each tree. In approximately 36% of the trees (133) the bark thickness at 0.15, 0.65, 1.30 m and at 2-m intervals thereafter, was measured. At each of these points, two rectangular-shaped pieces of bark, located 180 degrees one from the other, were extracted with a sharpened device designed for this task. An electronic calliper was used to measure the bark thickness with a precision of 0.1 mm. Volumes of individual log sections were estimated with the conoid formula (Equation 2.1) except for the stump that was assumed to be a cylinder. Total tree volumes were obtained by summing up all log sections.

$$V_i = \frac{0.00007854}{3} L\left(d_s^2 + d_l^2 + d_s d_l\right) \tag{2.1}$$

where V_i: log volume (m^3); L: log length (m); d_s: small end diameter (cm); d_l: large end diameter (cm).

Table 2.1 Summary of the data available.

	Eucalyptus grandis		*Pinus radiata*	
Statistic	whole dataset (ob)	under-bark subset	over-bark subset	whole dataset (ub)
N (trees)	374	133	428	674
n (measurements)	7917	1555	7850	10788
Mean *dbh* (cm)	21.0	22.1	37.8	34.7
Max. *dbh* (cm)	51.6	48.8	70.7	70.7
Min. *dbh* (cm)	6.0	6.3	5.1	4.0
Mean tree height (m)	22.2	23.0	25.1	23.5
Max. tree height (m)	39.4	39.4	40.8	40.8
Min. tree height (m)	6.3	6.7	3.3	2.0
Mean tree volume (m^3)	0.36	0.37	1.14	0.892
Max. tree volume (m^3)	3.24	2.52	4.99	4.99
Min. tree volume (m^3)	0.001	0.0098	0.0053	0.0035

Data for *P. radiata* were from the following forests located in the Auckland and Northland regions of New Zealand: Athenree, Maramuru, Tairua, Riverhead, Woodhill, Dargaville, Maharangi, and Whangarei (Dr. R.C. Woollons, personal data). From each forest 6 to 14 stands were sampled, with 9 trees on average sampled from each stand. For the first five forests both under-bark and over-bark diameters were measured, whereas for the last three forests only under-bark diameters were recorded. Measurement points for this species were located at 0.20, 0.50, 1.00, 1.40 and 2.00 for most trees, and thereafter at intervals of 1.50 m on average (minimum, 0.43; maximum, 4.20; standard deviation, 0.67 m). On average, 16 measures were taken from each *P. radiata* tree. Breast height (BH) is at 1.30 m in Uruguay and at 1.40 m in New Zealand.

Both datasets covered a wide range of *dbh* and height classes (Tables 2.2 and 2.3). A summary of both datasets by zone (*E. grandis*) and forest (*P. radiata*) is presented in Table 2.4.

Table 2.2 Number of trees by *dbh* and height classes from all zones (*E. grandis*)

dbh class (cm)	Height class (m)							
	<10	10-15	15-20	20-25	25-30	30-35	>35	Total
<10	13	5	1					19
10-15	5	63	31	2				101
15-20		5	60	47	5			117
20-25			20	42	16	1		79
25-30				10	8	8	2	28
30-35					3	7	8	18
35-40						1	6	7
40-45						1	1	2
>45						1	2	3
Total	18	73	112	101	32	19	19	374

Table 2.3 Number of trees by *dbh* and height classes from all forests (*P. radiata*)

dbh class (cm)	Height class (m)								Total
	<5	5-10	10-15	15-20	20-25	25-30	30-35	>35	
<10	34	27							61
10-15		51	9						60
15-20		34	40	4	1	1			80
20-25		3	31	19	7	8			68
25-30			4	26	14	20	4		68
30-35			4	16	20	29	9	1	79
35-40				6	14	32	12	5	69
40-45				1	10	18	22	4	55
45-50				1	7	16	20	13	57
50-55					2	4	17	14	37
55-60					2	3	11	7	23
>60						2	8	7	17
Total	34	115	88	73	77	133	103	51	674

In order to detect evident errors (such as typing errors), the profile of each tree was plotted and visually inspected. The quality of both datasets was excellent and only a few measurements (less than a dozen in both datasets) needed to be deleted, e.g. when there was an extreme deviation from the profile curve.

Table 2.4 Number of trees and sectional measurements by Zone (*E. grandis*) and forest (*P. radiata*)

Zone / Forest	Number of trees	Number of sectional measurements
7	162	3680
8	78	1520
9	134	2717
Total *E. grandis*	374	7917
Athenree	68	1301
Maramuru	83	1546
Tairua	82	1464
Riverhead	87	1673
Woodhill	108	1867
Dargaville	49	629
Maharangi	83	923
Whangarei	114	1383
Total *P. radiata*	674	10786

From each dataset 75% of the trees were used for model development (developmental subsets) and the remainder 25% were withheld for validation purposes (validation subsets). Taper datasets are intrinsically autocorrelated, implying that error variances are underestimated and hypothesis testing invalid. Therefore, using independent datasets to evaluate and validate the developed models is crucial. Most studies that set aside a subset of the data for validation perform the splitting randomly (Muhairwe, 1999; Figueiredo-Filho *et al.* 1996; Newnham, 1992; Perez *et al.* 1990; Byrne and Reed, 1986). However, Kozak and Smith (1993) suggested that the subset for estimating regression coefficients should be *selected* in a way that the sample covered the whole range of possible diameters and tree heights. In order to avoid sources of bias, the selection was initially performed at random. Then, the ranges of diameters and tree heights were compared to those in the original datasets and a few trees were interchanged to improve the coverage of developmental and validation datasets. Altogether, 506 and 280 trees formed the developmental subsets of *P. radiata* and *E. grandis* respectively, whereas 168 and 94 trees formed the validation subsets.

For developing models to predict *dub* from estimated *dob* (which will be further referred to as *bark models*), all data containing under-bark measurements were used in the case of *E. grandis* (i.e. 133 trees). In the case of *P. radiata*, a much larger dataset was available to fit the bark model (i.e. 428 trees). Therefore, this dataset was used to compare a number of sub-sampling strategies for developing over- and under-bark taper models

using the *composite* approach. That is, the bark model was fitted from various sub-samples extracted from the original data. Two sub-sampling methods were followed. In one of them, different proportions of *observations* were randomly selected from the original data, but all trees were included. In the other method, different proportions of *trees* were randomly selected. In the latter, all observations of selected trees were retained. For both sub-sampling methods the sampling intensities were 20, 40, 60, 80 and 100%. The selection of best bark models was firstly done using whole datasets.

2.2 METHODS AND PROCEDURES

2.2.1 Taper models evaluated

Taper models discussed in Sections 1.1.1, 1.1.2 and 1.1.3 were all considered for this study. They were fitted to three different data subsets, namely (i) *E. grandis*, over bark; (ii) *P. radiata*, under bark; and (iii) *P. radiata*, over bark. Some models were discarded at early stages due to various reasons outlined below. Because of its extremely simple formulation, Equation 1.6 (Kozak *et al.* 1969) gave very poor results in preliminary analyses and failed to converge several times when fitted by non-linear regression. Equations 1.8 and 1.9 were clearly inferior to the proposed variants (Equations 1.8a and 1.9a) for *E. grandis* and failed to converge when fitted to *P. radiata* data. Therefore, Equations 1.6, 1.8 and 1.9 were not further investigated.

Equation 1.12 performed exactly equal to Equation 1.11 in preliminary analyses. In fact, all fitting statistics as well as graphical plots of residuals were identical. Having the *same* model twice in the list of candidates would not contribute more information and may be misleading. As the model of Max and Burkhart (1976) is better known and is a pure taper model, it was retained and Equation 1.12 was removed from the list of candidates.

New variants specifically suited to the data were added to the list of candidate models. For instance, as residuals from Equation 1.11 showed an evident trend with *dbh*, the expression ln(*dbh*) was added to its third and fourth terms, which virtually eliminated this trend and improved all fit statistics. Other transformations such as *dbh* itself and other

ways of scaling *dbh* were also tried but with poorer results. Therefore, this variant of Max and Burkhart's model (Equation 2.11a) was further evaluated.

$$d = \sqrt{dbh^2\left[b_1\left(\frac{h}{H}-1\right)+b_2\left(\frac{h^2}{H^2}-1\right)+b_3\ln(dbh)\left(a_1-\frac{h}{H}\right)^2 I_1+b_4\ln(dbh)\left(a_2-\frac{h}{H}\right)^2 I_2\right]} \qquad (2.11a)$$

In the case of variable-exponent models, the selection of variables and transformations for the exponent term is crucial. An attempt was made to find the set of variables for the exponent term of Equation 1.14 that was best suited for the two species. For doing this the equation was transformed into a linear model by using logarithms and a large set of candidate variables was screened by a stepwise multiple regression procedure. The significance level imposed for a variable to enter or to stay in the model was p=0.0001 (this exigent level of significance was intended to compensate for the overestimation of the F values by virtue of the autocorrelated nature of the data and the quantum of the error degrees of freedom). Most variables selected differed between species, and the model comprising only variables found significant for both datasets was excessively simple and performed poorly compared to the original Kozak's (1988) model. Therefore, the variables for the exponent terms were defined individually for each subset (Equations 2.14a, 2.14b and 2.14c). In the case of *E. grandis* nine parameters (including p) were selected, which is the same number of parameters as in Kozak's model. In the case of *P. radiata*, however, 11 and 17 parameters were initially selected for the under- and over-bark subsets, respectively. In order to keep the models comparable, the number of parameters was restricted to nine. This was performed using the MAXR option instead of the STEPWISE option of Proc REG, the multiple regression procedure of the SAS System (SAS Institute Inc. 1989). The MAXR option adds one variable at a time, the entering variable at each step being the one making the largest contribution to R^2. After a variable is added, each variable in the regression is compared with each variable not in the regression to check which would make the largest contribution to R^2. When all comparisons have been made, MAXR switches to the two variables that result in the biggest net increase in R^2, and so forth. Before moving to the next step, the process is repeated until no further increase in R^2 can be obtained (Newnham 1992; SAS Institute Inc. 1989).

E. grandis - $$d=a_0 dbh^{a_1}\left(\frac{1-\sqrt{z}}{1-\sqrt{p}}\right)^{b_1\sqrt{z}+b_2\exp(z)+b_3/z+b_4(dbh/H)^2+b_5 dbh+b_6/H} \quad (2.14a)$$

P. radiata (dub)- $$d=a_0 dbh^{a_1} a_2^{dbh}\left(\frac{1-\sqrt{z}}{1-\sqrt{p}}\right)^{b_1\sqrt{z}+b_2(dbh/H)+b_3/h+b_4/h^2+b_5\left(H/\sqrt{h}\right)} \quad (2.14b)$$

P. radiata (dob)- $$d=a_0 dbh^{a_1} a_2^{dbh}\left(\frac{1-\sqrt{z}}{1-\sqrt{p}}\right)^{b_1 z^2+b_2 z^3+b_3\ln(z+0.001)+b_4(dbh/H)+b_5\left(H/\sqrt{h}\right)} \quad (2.14c)$$

For Muhairwe's (1999) model, Equation 1.17, the set of variables for the exponent term was also selected for each species separately through a multiple regression stepwise procedure. Even when a level of significance of p=0.0001 was imposed, 11 and 12 parameters were selected for *P. radiata* (for *dub* and *dob* respectively) and 13 for *E. grandis*. This was considered excessive and the MAXR option was invoked for the multiple regression procedure. The models were restricted to nine parameters (as in Equations 1.14, 2.14a, 2.14b, 2.14c and 1.15) and the following variables were selected:

E. grandis $$d=a_0 dbh^{a_1} a_2^{dbh}\left(1-\sqrt{z}\right)^{b_1\ln(z+0.001)+b_2\exp(z)+b_3(dbh/H)+b_4\ln(dbh)+b_5 H/\sqrt{h}+b_6(dbh/H)/h} \quad (2.17a)$$

P. radiata (dub) $$d = a_0 dbh^{a_1}\left(1-\sqrt{z}\right)^{b_0+b_1 z^2+b_2 z^3+b_3(dbh/H)+b_4/h+b_5/h^2+b_6/h^3} \quad (2.17b)$$

P. radiata (dob) $$d = a_0 dbh^{a_1}\left(1-\sqrt{z}\right)^{b_0+b_1 z+b_2(dbh/H)+b_3 H/\sqrt{h}+b_4(dbh/H)/h+b_5/h} \quad (2.17.c)$$

2.2.2 Criteria for model evaluation

The three most commonly used statistics for evaluating and comparing taper models found in published taper studies (Table 1.1) were computed on the developmental subsets to perform the first screening of candidate models. These statistics were the mean residual (B), the standard error of estimate (SEE) and the mean absolute deviation

(MAD) or mean absolute residual. The formulae for these statistics were displayed as Equations 1.18, 1.19 and 1.20.

A rank was assigned for each statistic and an overall rank was then computed by summing up all individual ranks. In the case of mean residual, its absolute value was considered for ranking purposes. For all three statistics a smaller value indicates a lower (better) score.

During the screening process, visual inspections of residual plots and analyses of residual statistics and normality tests yielded by Proc Univariate (SAS Institute Inc. 1990) were also performed. After the list of candidate models was reduced to a few models, the analysis of residual plots was a major consideration.

A few models that passed the screening stage were more intensively analysed with the validation datasets. These analyses included the evaluation of the projection ability of diameter along different parts of the stem, and the evaluation of merchantable height and total tree volume predictions. Merchantable height and volume predictions were analysed at two levels, overall and by *dbh* classes. Merchantable height was defined as the height at 90 percent of total tree height and it was symbolised as h_{90}. As sectional measurements were not explicitly taken at h_{90}, lineal interpolations between the nearest observations available for each tree needed to be carried out.

Estimations of diameters, merchantable heights and total tree volumes were obtained by applying the taper functions with parameters estimated from the relevant developmental datasets. The number of parameters in the taper equation was not considered in the calculation of SEEs for the validation subsets. The following formula, which is the same as the square root of the mean square error (RMSE) was used:

$$SEE_{validation} = RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n}}$$

The most accurate volume estimates can be obtained by numerical integration of the taper function. However, to be strictly comparable with observed volumes, estimated

volumes were calculated using the conoid formula for all sections but the lowest (which was assumed to be a cylinder). Some taper equations can be explicitly inverted to estimate heights but others cannot (e.g. variable exponent equations). When an explicit inversion is not feasible, numerical iterative methods such us the Newton-Raphson algorithm can be used. This was performed by using SAS's Proc Model with a SOLVE statement (Woodward, D. personal communication[3]).

A ranking procedure (largely based on the study by Kozak and Smith 1993) was used to compare the best models selected after the first screening. This procedure involved the calculation of a global rank resulting from the summation of the following ranks:

(i) overall ranks of diameter, volume and merchantable height (one overall rank for each statistic);
(ii) sum of the ranks for diameter by tree position (one sum for each statistic);
(iii) sum of the ranks for volume by dbh classes (one sum for each statistic); and
(iv) sum of the ranks for merchantable height by dbh classes (one sum for each statistic).

In addition to this ranking procedure (further discussed in Chapter 3), plots of residuals and statistics describing the distribution of the residuals were carefully examined. The level of significance of the parameters in the final models was checked by analysing the models as *mixed models* (i.e. taking into account the autocorrelated nature of the data).

2.2.3 Modelling bark thickness

A polynomial of the type of Equation 1.23 (Section 1.4) with candidate terms up to the 10^{th} power, was fitted to both datasets. The selection of terms was performed with a stepwise linear regression procedure in which the dependent variable was the ratio dub^2/dob^2. The level of significance for a variable to enter or to stay in the model was set at $p<0.0001$. Once significant variables were selected, the models were re-arranged to estimate *dub* directly and the parameters were refined by non-linear regression. The intercept and the terms in y, y^2 and y^{10} were found significant for both species (Equations 2.23a and 2.23b). For *E. grandis*, the term in y^9 was also significant.

[3] Woodward, Donna. SAS Institute Inc. Technical Support, Area Statistics

E. grandis $$dub = \sqrt{dob^2\left(a_0 + a_1 y + a_2 y^2 + a_3 y^9 + a_4 y^{10}\right)} \quad (2.23a)$$

P. radiata $$dub = \sqrt{dob^2\left(a_0 + a_1 y + a_2 y^2 + a_3 y^{10}\right)} \quad (2.23b)$$

The plots in Figure 1.2 (Section 1.4) indicated that the ratio of inside-bark sectional area to outside-bark sectional area for *P. radiata* did not follow a pattern as clear as it did for *E. grandis*. Therefore, a simple polynomial on *y* may not be sufficient to describe the curve displayed in Figure 1.2 for *P. radiata*. A multiple regression analysis based on a polynomial on *y*, but which also included other variables representing tree size (H, dbh, ln(H), ln(dbh), H*dbh), total tree taper (dbh/H) and other variables, needed to be evaluated. The inverse of height and the inverse of squared height were considered valuable terms. As these terms adopt very large values at the lower part of the tree and small values at the top, they have an important influence on the bottom section, which showed the largest variation. The selection of candidate variables was carried out by stepwise linear regressions with a significance level of $p<0.0001$. Eight and twelve parameters (including the intercept) for *E. grandis* and *P. radiata* respectively satisfied the imposed significance level (Equations 2.24a and 2.24b). In addition to these models, simplified versions containing only 7 parameters (Models 2.25a and 2.25b), as Gordon's (1983) model were developed using the MAXR option.

E. grandis $$\frac{dub^2}{dob^2} = a_0 + a_1 y^{10} + a_2 \ln(H) + \frac{a_3}{h} + \frac{a_4}{h^2} + \frac{a_5}{h^3} + \frac{a_6}{y} + a_7\sqrt{y} \quad (2.24a)$$

P. radiata $$\frac{dub^2}{dob^2} = a_0 + a_1 y^{10} + a_2 \ln(H) + a_3 \ln(dbh) + a_4 \frac{dbh}{H} + \frac{a_5}{h} \quad (2.24b)$$

$$+ \frac{a_6}{h^2} + \frac{a_7}{h^3} + \frac{a_8}{y} + a_9\sqrt{y} + a_{10}\frac{H}{\sqrt{h}} + a_{11}\frac{dbh/H}{h}$$

E. grandis $$\frac{dub^2}{dob^2}=a_0+a_1y+a_2\ln(H)+\frac{a_3}{h}+\frac{a_4}{h^2}+\frac{a_5}{h^3}+a_6\sqrt{y}$$ (2.25a)

P. radiata $$\frac{dub^2}{dob^2}=a_0+a_1y^8+a_2\ln(dbh)+a_3\ln(H)+a_4\sqrt{y}+a_5\frac{H}{\sqrt{h}}+a_6\frac{dbh/H}{h}$$ (2.25b)

Equations 2.23a and 2.23b, Models 2.24a, 2.24b, 2.25a and 2.25b were re-fitted as non-linear models predicting *dub* directly. All fit statistics and parameter estimates presented in Chapter 3 refer to non-linear versions with *dub* as the independent variable.

The inclusion of age terms was also tried and some of these terms were statistically significant, despite overall indicators of model performance not changing substantially. As information on age is not always available in practice, age terms were not included.

The selection of the best bark models was based on fit statistics (e.g. B, SEE, MAD) and plots of residuals. The level of significance of the parameters was checked by analysing the equations as mixed models.

CHAPTER 3

DEVELOPING AND EVALUATING TAPER MODELS

3.1 SCREENING OF CANDIDATE MODELS

All candidate taper models were fitted to developmental subsets and selected fit statistics (B, SEE and MAD) were computed (Table 3.1). Rankings for each statistic and overall ranks are shown in Table 3.2. The decision of the number of models to select for further evaluation was not straightforward. There was not a well-defined limit separating the best models from the worst models, although the models here defined as *simple* (Section 1.1.1.1) clearly occurred in the latter stratum.

Variable-exponent models consistently outperformed other categories. A second stratum was occupied by high degree polynomials such as the variants of the models proposed by Bruce *et al.* (1968), Hilt (1980) and Real and Moore (1988). The segmented group ranked third, with the modified version of the Max and Burkhart (1976) model (i.e. Equation 2.11a) performing consistently better than other segmented models. All proposed variants (i.e. model number followed by *a* or *b*) performed better than their original formulations for all datasets.

After careful examination of the rankings it was decided to retain five models for further examination. The decision was not based purely on the rankings, but other considerations were also taken into account as explained below. The selected models were:

1. Equations 2.14a 2.14b and 2.14c, based on the model by Kozak (1988) but with independent variables specifically selected for each subset
2. Equations 2.17a, 2.17b and 2.17c, based in the model by Muhaiwe (1999)
3. Equation 1.16, a variable exponent model which independent variables were specifically selected to avoid multicollinearity between independent variables
4. Equation 1.8a, a variant of the polynomial-based model described by Bruce at *al.* (1968)
5. Equation 2.11a, a variant of the classic segmented taper model introduced by Max and Burkhart (1976)

The original variable exponent model (Equation 1.14) was not selected because it was very similar to its variants (Equations 2.14a, 2.14b and 2.14c), which were considered sound surrogates for that model. Equation 1.15 performed very well, ranking fourth for *E. grandis*, second for the under-bark data of *P. radiata*, and fifth for the over-bark data of *P. radiata*. However, the variable-exponent group was already represented by three models. Equation 1.9a performed better than Equation 2.11a for *E. grandis*, and Equation 1.10a performed slightly better than Equation 2.11a for *P. radiata*. However, these two polynomial-based models were not as consistent as Equation 2.11a and a better model of their category was already chosen (Equation 1.8a). Furthermore, the inclusion of a segmented model as Equation 2.11a was considered of much interest, as this model can be explicitly integrated providing tree volume and volume ratio equations (Byrne and Reed 1986)

Table 3.1 Fit statistics for all taper models (developmental subsets).

Equation	*E. grandis* (dob)			*P. radiata* (dub)			*P. radiata* (dob)		
	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE
1.1	0.0934	0.7723	1.2409	0.1959	1.3265	1.8598	0.2623	1.6433	2.2706
1.2	-0.0168	0.7568	1.1867	0.0078	1.2410	1.7613	0.0843	1.5510	2.1851
1.3	0.0273	0.7604	1.2034	0.0413	1.2347	1.7664	0.1208	1.5346	2.1891
1.3a	0.0093	0.7531	1.1923	0.0334	1.2298	1.7598	0.1060	1.5293	2.1800
1.4	0.0847	0.7713	1.2114	0.0751	1.1552	1.6318	0.1635	1.4790	2.1257
1.4a	0.0859	0.7696	1.2114	0.0760	1.1536	1.6308	0.1621	1.5270	2.1255
1.5	0.0544	0.7377	1.0327	0.1468	1.5353	2.0059	0.1132	1.8611	2.4460
1.7	0.0873	0.6522	0.9053	0.2070	1.2290	1.6510	0.2259	1.2897	1.7157
1.8a	0.0002	0.5176	0.7567	-0.0311	1.0132	1.4849	0.0200	1.1092	1.6697
1.9a	0.0004	0.5246	0.7742	0.0584	1.2044	1.7403	0.0345	1.2155	1.8866
1.1	0.1805	0.6381	0.9967	-0.9707	1.8062	2.4949	0.2387	1.1561	1.7550
1.10a	0.0223	0.6250	0.8995	0.0140	1.0972	1.5813	0.0292	1.1685	1.6951
1.11	0.0909	0.6111	0.8572	0.2108	1.2059	1.6805	0.2298	1.2739	1.7452
2.11a	-0.0230	0.5391	0.7715	0.0674	1.1060	1.5661	0.0879	1.1639	1.6474
1.13	-0.0254	0.5967	0.8609	0.0680	1.1539	1.6161	0.0989	1.2234	1.7102
1.14	-0.0305	0.5288	0.7607	-0.0268	0.9705	1.4146	-0.0042	1.0065	1.4987
2.14a,b,c	-0.0016	0.4806	0.6958	-0.0289	0.9698	1.4114	-0.0047	0.9902	1.4782
1.15	-0.0121	0.5155	0.7377	0.0035	1.0093	1.4540	0.0134	1.0575	1.5360
1.16	-0.0250	0.5281	0.7627	-0.0207	0.9950	1.4553	0.0090	1.0205	1.5340
2.17a,b,c	-0.0021	0.4913	0.7070	-0.0213	0.9321	1.3640	-0.0160	1.0008	1.4823

Table 3.2 Model rankings for each statistic and overall rankings

Equation	*E. grandis* (dob)				*P. radiata* (dub)				*P. radiata* (dob)			
	B	MAD	SEE	overall	B	MAD	SEE	overall	B	MAD	SEE	overall
1.1	19	20	20	20	17	18	18	18	20	19	19	20
1.2	7	16	15	13	2	17	16	11	9	18	17	14
1.3	12	17	17	17	10	16	17	15	14	17	18	18
1.3a	5	15	16	11	9	15	15	14	12	16	16	15
1.4	15	19	19	19	14	11	11	12	16	14	15	17
1.4a	16	18	18	18	15	9	10	10	15	15	14	16
1.5	14	14	14	14	16	19	19	19	13	20	20	19
1.7	17	13	12	15	18	14	12	16	17	13	10	12
1.8a	**1**	**4**	**4**	**3**	**8**	**6**	**6**	**7**	**6**	**6**	**7**	**6**
1.9a	2	5	8	5	11	12	14	13	8	10	13	9
1.1	20	12	13	16	20	20	20	20	19	7	12	11
1.10a	8	11	11	9	3	7	8	6	7	9	8	7
1.11	18	10	9	12	19	13	13	17	18	12	11	13
2.11a	**9**	**8**	**7**	**7**	**12**	**8**	**7**	**8**	**10**	**8**	**6**	**8**
1.13	11	9	10	10	13	10	9	9	11	11	9	10
1.14	13	7	5	8	6	3	3	4	1	3	3	2
2.14a,b,c	**3**	**1**	**1**	**1**	**7**	**2**	**2**	**3**	**2**	**1**	**1**	**1**
1.15	6	3	3	4	1	5	4	2	4	5	5	5
1.16	**10**	**6**	**6**	**6**	**4**	**4**	**5**	**5**	**3**	**4**	**4**	**4**
2.17a,b,c	**4**	**2**	**2**	**2**	**5**	**1**	**1**	**1**	**5**	**2**	**2**	**3**

Note: Models selected for further analyses are highlighted with bold letters.

With regard to the statistics used for screening the models, the B statistic (mean residual) behaved more independently than did the other two statistics. In the case of *E. grandis*, for example, Equation 1.3a was ranked fifth by the B statistic but it was very poorly ranked by the MAD and SEE (15^{th}, 16^{th} positions respectively). A similar situation was evident for *P. radiata* (*dub*) where Equation 1.2 was ranked second by the B statistic and 17^{th} and 16^{th} by the MAD and SEE. The reason for that behaviour of the B statistic is that positive and negative residuals may cancel each other out even with a very imprecise model. Therefore, this criterion should not be used alone, especially when considered at the overall level, to judge and compare models.

The MAD showed a similar pattern to the SEE, the main difference being that the former does not magnify the weight of the largest residuals as the latter does. This feature of the MAD makes it more comparable with the commonly used mean residual (B) with the advantage that the residuals cannot be cancelled out. The MAD statistic, which can be interpreted as the average error of prediction regardless of sign, was then kept as a ranking criterion along with B and SEE for evaluating and comparing the five models against validation subsets.

Diameter residuals for selected models showed no clear patterns when plotted against predicted values, percentage height, *dbh* and tree height. Parameter estimates and asymptotic standard errors of selected models are shown in Table 3.3.

Table 3.3 Parameter estimates and asymptotic standard errors (ASE) of selected taper models.

		E. grandis (dob)		*P. radiata* (dub)		*P. radiata* (dob)	
Equation	Parameter *	Estimate	ASE	Estimate	ASE	Estimate	ASE
1.8a	b0	0.0019137	0.00016	-0.0003116	0.00014	0.0016414	0.00019
	b1	0.9532228	0.00406	0.8184187	0.00383	0.9428256	0.00532
	b2	-0.0079081	0.00096	-0.0110262	0.00049	-0.0073425	0.00066
	b3	0.0262282	0.00113	0.0272013	0.00096	0.0180073	0.00128
	b4	-0.0000894	0.00001	-0.0000385	0.00000	-0.0000656	0.00000
	b5	0.0087623	0.00049	0.0001981	0.00002	0.0005809	0.00002
	b6	-0.0001145	0.00001	0.0000040	0.00000	0.0000040	0.00000
	c	1.6714036	0.00791	1.7000000	0.01013	1.7435939	0.01448
2.11a	a1	0.7593559	0.00405	0.7428539	0.00556	0.7960876	0.00599
	a2	0.0631599	0.00120	0.0696211	0.00234	0.0963214	0.00235
	b1	-3.5149098	0.03329	-2.9438636	0.03621	-3.4374623	0.05227
	b2	1.6666338	0.01856	1.4483698	0.02007	1.6932176	0.02850
	b3	-0.4521498	0.00701	-0.3173639	0.00618	-0.3187316	0.00823
	b4	48.9365393	2.26508	13.6676547	1.06509	11.1366342	0.60451
2.14a,b,c	p	0.2688427	0.00885	0.1787144	0.00685	0.1859999	0.00940
	a0	0.9865925	0.01210	1.3645542	0.03489	1.5686034	0.04573
	a1	0.9382299	0.00196	0.8228798	0.00880	0.7928302	0.00984
	b1 / a2	0.0011918	0.00005	1.0030747	0.00024	1.0033704	0.00025
	b2 / b1	-0.4787293	0.02432	0.4885155	0.00653	1.4686888	0.04835
	b3 / b2	0.4017619	0.01050	0.1411555	0.00360	-0.9240403	0.04947
	b4 / b3	0.0593946	0.00502	-0.0026583	0.00028	-0.1607549	0.00621
	b5/ b4	-0.0007792	0.00020	0.1393396	0.00776	0.1167483	0.00455
	b6 / b5	2.2738610	0.10301	-0.0094525	0.00074	-0.0041751	0.00030
1.16	a0	0.9119130	0.00608	0.8804301	0.00705	0.9481726	0.00921
	a1	0.9704829	0.00506	1.0180166	0.00426	0.9895282	0.00464
	a2	0.0641536	0.00557	-0.0120814	0.00429	0.0284921	0.00477
	b1	0.5665675	0.00970	0.7659725	0.00866	0.6857261	0.00936
	b2	0.4185130	0.00896	0.4279568	0.01148	0.4396912	0.01251
	b3	-0.1562910	0.02261	-0.3261356	0.02427	-0.0559025	0.02600
	b4	-0.1898582	0.02002	-0.7020145	0.02185	-0.4690056	0.02517
	b5	0.0154833	0.00061	0.0047966	0.00033	0.0033150	0.00034
2.17a,b,c	a0	1.8617844	0.05044	0.6400203	0.01301	0.5776916	0.01619
	a1	0.9358676	0.00865	1.0631322	0.00399	1.1346308	0.00836
	a2 / b0	1.0011254	0.00031	-0.0205954	0.01102	-0.1209232	0.01810
	b1	-0.3615384	0.01057	0.8302343	0.03297	0.4837346	0.01109
	b2	0.3197242	0.00380	-0.3539122	0.03168	0.1358230	0.00543
	b3	0.2210044	0.01178	0.1406034	0.00373	0.0184053	0.00231
	b4	-0.1458189	0.00695	-0.8669686	0.02739	0.2316543	0.01755
	b5	0.0285517	0.00322	0.0852768	0.00564	-1.2971317	0.04504
	b6	-0.1489390	0.01249	-0.0048365	0.00045	0.0102172	0.00053

* When different, parameters before the / correspond to *E. grandis* and after the /, to *P. radiata*.

3.2 FURTHER EVALUATION OF SELECTED TAPER MODELS

With parameter estimates obtained by fitting the models to developmental datasets, selected models were further evaluated and compared against validation subsets. The same statistics used for model screening (B, MAD and SEE) were calculated for diameter (Table 3.4), merchantable height (Table 3.5) and total tree volume (Table 3.6).

Using the results in Tables 3.4 to 3.6, the five taper estimation systems were compared through a detailed ranking procedure (Table 3.7). The models' performances were evaluated for three attributes: diameter, merchantable height and total tree volume. Due to the importance of all three attributes in practice, they were equally weighted.

The first three rows of Table 3.7 show the sum of ranks for overall bias, MAD and SEE, respectively, for diameter, merchantable height (h_{90}) and total tree volume. For example, Equation 2.14a (*E. grandis*) had the lowest bias (rank 1) for diameter, merchantable height and total tree volume. Therefore, it totalled a rank of 3 in the first row, which is the lowest value in this row for this species and thus corresponds to a relative rank of (1). As discussed earlier, a proper evaluation and validation of taper estimation systems should not be limited to analysing their *overall* performance. It is also necessary to evaluate their prediction ability for diameters at different positions along the stem. Likewise, their performance for merchantable height and total tree volume estimations should be tested by *dbh* classes (Kozak and Smith 1993). Sums of the ranks in rows 4-12 of Table 3.7 were generated in a way that each taper estimation system was assigned a rank separately for every percent height, for diameter, and for every *dbh* class for merchantable height and volume. These ranks were computed for each statistic (B, MAD and SEE). Finally, totals were calculated in row 13 of Table 3.7 by simply summing the ranks of all 12 rows and a final rank (row 14) was assigned to each taper estimation system.

Table 3.4 Biases (B), MAD's and SEE's of diameter from ground to top (validation subsets)

Relative height	n	Equation 1.8a			Equation 2.11a			Equations 2.14a,b,c			Equation 1.16			Equations 2.17a,b,c		
		B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE
E. grandis (over bark)																
0 - 5%	184	0.2608	0.8115	1.1871	0.1887	0.8919	1.2506	0.1026	0.7549	1.0654	0.1819	0.9188	1.2536	-0.0542	0.7905	1.1071
5 - 10%	115	0.1987	0.3073	0.4864	0.1878	0.3867	0.5355	-0.2020	0.3412	0.4494	-0.2411	0.3795	0.5477	-0.1819	0.3410	0.4637
10 - 20%	184	-0.0704	0.4220	0.5729	-0.2181	0.4483	0.5975	-0.0285	0.4020	0.5341	0.0434	0.4231	0.5680	0.0525	0.4141	0.5540
20 - 30%	189	-0.2248	0.4610	0.5982	-0.2177	0.4454	0.5720	0.0816	0.4101	0.5354	0.1836	0.4364	0.5602	0.1208	0.4428	0.5623
30 - 40%	188	-0.2228	0.4468	0.6145	-0.1411	0.4175	0.5607	0.0142	0.4250	0.5627	0.0877	0.4244	0.5653	0.0195	0.4385	0.5865
40 - 50%	187	-0.1276	0.4281	0.5461	-0.0674	0.4014	0.5075	-0.0722	0.4337	0.5425	-0.0814	0.4202	0.5234	-0.0754	0.4455	0.5645
50 - 60%	189	-0.0042	0.4370	0.6112	-0.0295	0.4456	0.5954	-0.1243	0.4524	0.6347	-0.2500	0.4896	0.6487	-0.1174	0.4505	0.6417
60 - 70%	182	0.0990	0.5164	0.6879	-0.0087	0.5353	0.6944	-0.0877	0.4930	0.6748	-0.2926	0.5655	0.7319	-0.0734	0.4892	0.6773
70 - 80%	186	0.1465	0.6281	0.8096	0.0789	0.6583	0.8445	0.0491	0.5520	0.7359	-0.1611	0.6438	0.8321	0.0460	0.5637	0.7468
80 - 90%	181	-0.1114	0.6132	0.7765	-0.0673	0.6649	0.8246	-0.0306	0.5324	0.6983	-0.1044	0.6316	0.7891	-0.0710	0.5506	0.7169
90 - 100%	208	-0.1317	0.2492	0.4224	-0.1696	0.2675	0.4564	-0.0370	0.2118	0.3591	-0.0103	0.2114	0.3686	-0.0642	0.2197	0.3710
Total/average	1993	-0.0268	0.4866	0.6963	-0.0522	0.5062	0.5064	-0.0244	0.4561	0.6446	-0.0511	0.5042	0.7085	-0.0311	0.4690	0.6635
P. radiata (under bark)																
0 - 5%	372	0.4390	1.1586	1.5474	0.4231	1.3726	1.7966	0.3850	1.1060	1.4123	0.3527	1.1630	1.4950	0.1949	0.9361	1.2327
5 - 10%	242	0.0472	0.7560	0.9679	0.4167	1.1026	1.3854	0.0953	0.7754	0.9776	0.0499	0.7599	0.9571	0.3316	0.7787	0.9843
10 - 20%	311	-0.0879	0.8508	1.1308	0.0571	0.9777	1.2322	-0.2204	0.8878	1.1686	0.1429	0.8382	1.0965	0.2360	0.8734	1.1353
20 - 30%	228	-0.0326	0.9523	1.2493	0.1199	0.9651	1.2089	-0.0960	0.9472	1.1967	0.3171	0.9859	1.3018	0.0993	0.9214	1.2304
30 - 40%	223	0.0295	1.0539	1.4027	0.1568	1.0112	1.3007	0.1243	1.0147	1.3149	0.2085	1.0787	1.4263	-0.0426	1.0312	1.3624
40 - 50%	223	0.0610	1.1455	1.4756	0.0990	1.1187	1.4635	0.2491	1.1087	1.4285	-0.0577	1.1682	1.4993	-0.1136	1.1057	1.4065
50 - 60%	203	-0.0160	1.1895	1.5045	-0.0729	1.2637	1.6666	0.1803	1.1368	1.4745	-0.3947	1.2409	1.5504	-0.1462	1.1560	1.4456
60 - 70%	200	-0.1353	1.2399	1.5979	-0.2689	1.4914	1.9112	0.0251	1.1473	1.4853	-0.5555	1.2719	1.6104	-0.0973	1.1738	1.4919
70 - 80%	201	-0.1816	1.3010	1.7135	-0.0750	1.5121	1.9846	-0.1782	1.1771	1.5605	-0.3844	1.2474	1.6190	-0.0353	1.2104	1.5828
80 - 90%	193	-0.1914	1.0002	1.3368	0.1619	1.1491	1.5493	-0.4108	0.9110	1.2293	0.0108	0.8521	1.1297	-0.0652	0.8495	1.1349
90 - 100%	266	0.0603	0.2945	0.6445	0.1542	0.3265	0.6974	-0.1194	0.3219	0.6580	0.1793	0.2768	0.6270	0.0052	0.2572	0.5714
Total/average	2662	0.0271	0.9803	1.3449	0.1311	1.1062	1.5021	0.0203	0.9480	1.2776	0.0281	0.9763	1.3222	0.0551	0.9169	1.2426
P. radiata (over bark)																
0 - 5%	314	0.3848	1.2408	1.8451	0.3627	1.2350	1.5918	0.1079	0.8667	1.2357	0.2727	0.9477	1.2839	0.2288	0.9044	1.1977
5 - 10%	184	0.0685	0.7501	0.9785	0.6047	1.0841	1.3759	0.2867	0.7155	0.9205	-0.1159	0.6172	0.8702	0.2164	0.7029	0.8904
10 - 20%	255	-0.3918	0.9197	1.2698	-0.0567	1.0108	1.3153	0.1260	0.8799	1.1465	-0.1396	0.8342	1.1705	-0.1471	0.8277	1.1361
20 - 30%	190	-0.3270	1.0290	1.3279	-0.0694	1.0464	1.2824	0.0024	0.9899	1.2713	0.2432	0.9974	1.3285	-0.0886	0.9430	1.2231
30 - 40%	179	-0.1103	1.1014	1.4373	0.0691	1.0649	1.3510	-0.2033	1.0849	1.3891	0.2533	1.1111	1.4727	0.0503	1.0539	1.3600
40 - 50%	171	0.1440	1.2170	1.5619	0.1811	1.1742	1.5275	-0.2364	1.1884	1.4887	0.0648	1.2058	1.5637	0.1922	1.1556	1.4727
50 - 60%	153	0.2420	1.2487	1.6069	0.1344	1.2673	1.6928	-0.1645	1.1782	1.4955	-0.2538	1.2241	1.5825	0.1952	1.1829	1.5025
60 - 70%	149	0.0801	1.2993	1.7054	-0.1947	1.4601	1.9312	-0.1032	1.1808	1.5440	-0.5610	1.3061	1.7193	0.0054	1.2049	1.5546
70 - 80%	146	-0.0519	1.3219	1.7779	-0.2553	1.5131	2.0226	-0.0519	1.1851	1.5994	-0.4582	1.2937	1.7374	-0.2095	1.2303	1.6209
80 - 90%	145	-0.2074	1.1184	1.4581	0.0049	1.2376	1.6010	-0.1196	0.9056	1.1952	-0.0229	0.9439	1.2403	-0.4120	0.9456	1.2524
90 - 100%	194	0.0046	0.3833	0.7698	0.1791	0.3944	0.7926	-0.0146	0.3330	0.6762	0.2326	0.3540	0.7398	-0.1093	0.3356	0.6759
Total/average	2080	-0.0053	1.0471	1.4700	0.1109	1.1171	1.5045	-0.0125	0.9321	1.2726	-0.0078	0.9590	1.3412	0.0088	0.9289	1.2628

Table 3.5 Biases (B), MAD's and SEE's of merchantable height by *dbh* classes (validation subsets)

dbh class	n	Equation 1.8a			Equation 2.11a			Equations 2.14a,b,c			Equation 1.16			Equations 2.17a,b,c		
		B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE
E. grandis (over bark)																
< 10 cm	5	-0.1511	0.1653	0.2187	-0.2013	0.2013	0.2661	0.1573	0.2234	0.2647	-0.1913	0.1913	0.2431	0.0145	0.1564	0.1865
10 - 15 cm	20	-0.2051	0.2598	0.3132	-0.2698	0.3184	0.3720	0.0006	0.1728	0.2081	-0.2301	0.2653	0.3071	-0.0850	0.1776	0.2319
15 - 20 cm	32	-0.1529	0.3494	0.4002	-0.1837	0.4045	0.4649	-0.0512	0.2625	0.3011	-0.1150	0.3213	0.3648	-0.0864	0.2923	0.3316
20 - 25 cm	25	-0.0924	0.3608	0.4461	-0.0943	0.3989	0.5010	0.0046	0.3168	0.3881	0.0347	0.3441	0.4207	-0.0254	0.3300	0.4006
25 - 30 cm	7	-0.2800	0.4098	0.5317	-0.2409	0.4955	0.5976	-0.1928	0.3267	0.4332	-0.0394	0.4026	0.4832	-0.2178	0.3465	0.4574
30 - 35 cm	2	0.2441	0.3068	0.3921	0.4713	0.4713	0.5206	0.2466	0.3451	0.4242	0.6489	0.6489	0.6888	0.2444	0.3307	0.4113
> 35 cm	3	-0.7052	0.7052	0.7947	-0.5507	0.5507	0.6587	-0.6167	0.6167	0.7306	-0.1467	0.3170	0.3934	-0.6326	0.6326	0.7408
Total/average	94	-0.1664	0.3385	0.4187	-0.1812	0.3867	0.4688	-0.0365	0.2736	0.3445	-0.0829	0.3214	0.3847	-0.0847	0.2864	0.3607
P. radiata (under bark)																
< 10 cm	12	0.0181	0.0942	0.1188	0.1628	0.1750	0.2123	0.0081	0.0982	0.1256	0.0031	0.0920	0.1102	-0.0993	0.1269	0.1572
10 - 15 cm	16	0.0102	0.1421	0.2305	0.1282	0.1804	0.2839	-0.0983	0.1596	0.1948	-0.0177	0.1230	0.1620	-0.1100	0.1828	0.2149
15 - 20 cm	22	0.0468	0.1733	0.2473	0.1542	0.2053	0.3118	-0.0319	0.1327	0.1886	0.0835	0.1386	0.2025	-0.0054	0.1460	0.2041
20 - 25 cm	17	0.0558	0.2267	0.2599	0.2654	0.3500	0.4509	-0.2647	0.2648	0.3681	-0.0434	0.1743	0.2616	-0.1194	0.1727	0.2498
25 - 30 cm	18	0.4838	0.7889	1.0463	0.7965	1.0099	1.3548	0.0881	0.6697	0.8388	0.3172	0.6658	0.8226	0.3070	0.7469	0.9454
30 - 35 cm	21	0.1970	0.4989	0.7202	0.4959	0.7034	0.9791	-0.1315	0.5417	0.6644	0.1977	0.4682	0.6336	0.0885	0.5014	0.6871
35 - 40 cm	17	0.1014	0.4784	0.7301	0.4458	0.6684	0.9489	-0.2434	0.5788	0.7669	0.1834	0.4456	0.6900	0.0078	0.4836	0.7362
40 - 45 cm	9	-0.1909	0.3731	0.4368	0.2019	0.3690	0.6377	-0.5556	0.5996	0.6513	0.0082	0.2248	0.3259	-0.2675	0.4090	0.4591
45 - 50 cm	17	-0.3326	0.4603	0.5312	-0.0902	0.4189	0.4829	-0.5429	0.6462	0.7393	0.0224	0.3473	0.4398	-0.2707	0.4651	0.5661
50 - 55 cm	10	-0.3394	0.5931	0.6607	-0.1321	0.5155	0.6446	-0.5191	0.7189	0.7806	0.1171	0.4820	0.6045	-0.2179	0.5573	0.6436
55 -60 cm	4	-0.6139	0.6139	0.6404	-0.6597	0.6597	0.6979	-0.5113	0.5113	0.5936	-0.1677	0.1677	0.2456	-0.2725	0.3506	0.3936
> 60 cm	5	-0.6608	0.6608	0.8160	-0.6539	0.6539	0.7804	-0.7529	0.7537	0.9404	-0.0877	0.4616	0.5143	-0.4098	0.5894	0.7190
Total/average	168	0.0046	0.3974	0.5928	0.2254	0.4831	0.7427	-0.2209	0.4404	0.6024	0.0872	0.3246	0.4942	-0.0582	0.3841	0.5623
P. radiata (over bark)																
< 10 cm	9	0.0708	0.1044	0.1434	0.1948	0.1948	0.2334	0.0268	0.0976	0.1327	0.0342	0.0915	0.1197	-0.1770	0.1970	0.2169
10 - 15 cm	12	0.1413	0.1556	0.2690	0.2291	0.2291	0.3237	-0.0175	0.1228	0.1718	0.0624	0.1125	0.1744	-0.2016	0.2393	0.2653
15 - 20 cm	10	0.3558	0.3558	0.4561	0.4313	0.4313	0.5208	0.2070	0.2218	0.3030	0.2898	0.2898	0.3602	0.0212	0.1707	0.2216
20 - 25 cm	10	0.1817	0.3170	0.3610	0.2958	0.3884	0.4637	-0.0286	0.1775	0.1986	0.1293	0.2224	0.2487	-0.2158	0.2158	0.2876
25 - 30 cm	9	0.2672	0.7467	0.9888	0.4281	0.7835	1.1198	0.0674	0.6517	0.8316	0.1946	0.6329	0.7756	-0.0862	0.6247	0.8555
30 - 35 cm	17	0.2389	0.5394	0.6962	0.4492	0.6611	0.8285	0.0800	0.4825	0.6223	0.2632	0.4988	0.6215	-0.0955	0.4706	0.6083
35 - 40 cm	9	-0.0636	0.2718	0.3102	0.1574	0.3957	0.4357	-0.1094	0.2633	0.3064	0.1049	0.2351	0.2985	-0.2789	0.3163	0.3828
40 - 45 cm	7	-0.3482	0.3482	0.4039	-0.0250	0.2061	0.2361	-0.4153	0.4153	0.4880	-0.0733	0.1845	0.2346	-0.5702	0.5702	0.6166
45 - 50 cm	16	-0.3182	0.4613	0.5050	-0.0324	0.3741	0.4245	-0.2493	0.4500	0.5057	0.0551	0.3020	0.3789	-0.3821	0.5043	0.5702
50 - 55 cm	10	-0.3901	0.6412	0.7147	-0.1046	0.5143	0.6241	-0.2585	0.5858	0.6686	0.0789	0.4900	0.5891	-0.3660	0.6241	0.6958
55 -60 cm	4	-0.7505	0.7505	0.7816	-0.6497	0.6497	0.6960	-0.4223	0.4223	0.5071	-0.2997	0.2997	0.3698	-0.4819	0.4819	0.5690
> 60 cm	5	-0.8836	0.8836	0.9933	-0.6673	0.6673	0.7750	-0.6076	0.6583	0.8033	-0.2574	0.4401	0.5114	-0.6307	0.6852	0.8199
Total/average	118	-0.0389	0.4380	0.5847	0.1512	0.4507	0.6091	-0.0937	0.3738	0.5079	0.0985	0.3260	0.4482	-0.2493	0.4147	0.5409

Table 3.6 Biases (B), MAD's and SEE's of total tree volume by *dbh* classes (validation subsets)

dbh class	n	Equation 1.8a			Equation 2.11a			Equations 2.14a,b,c			Equation 1.16			Equations 2.17a,b,c		
		B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE	B	MAD	SEE
E. grandis (over bark)																
< 10 cm	5	0.0007	0.0008	0.0010	0.0010	0.0013	0.0015	0.0017	0.0017	0.0020	0.0003	0.0008	0.0009	0.0018	0.0018	0.0021
10 - 15 cm	20	0.0012	0.0028	0.0041	-0.0006	0.0025	0.0039	0.0009	0.0026	0.0039	-0.0003	0.0026	0.0038	0.0013	0.0027	0.0040
15 - 20 cm	32	-0.0004	0.0094	0.0117	-0.0036	0.0097	0.0118	-0.0014	0.0092	0.0113	-0.0019	0.0092	0.0114	-0.0009	0.0092	0.0113
20 - 25 cm	25	-0.0001	0.0129	0.0156	0.0001	0.0129	0.0155	0.0021	0.0135	0.0161	0.0026	0.0138	0.0162	0.0024	0.0140	0.0165
25 - 30 cm	7	-0.0057	0.0198	0.0227	0.0008	0.0206	0.0236	0.0003	0.0189	0.0225	0.0018	0.0183	0.0224	-0.0006	0.0182	0.0226
30 - 35 cm	2	0.0060	0.0305	0.0311	0.0221	0.0312	0.0382	0.0125	0.0277	0.0304	0.0108	0.0212	0.0238	0.0062	0.0253	0.0260
> 35 cm	3	-0.0619	0.0661	0.0996	-0.0454	0.0573	0.0892	-0.0690	0.0690	0.1031	-0.0633	0.0633	0.1010	-0.0802	0.0802	0.1100
Total/average	94	-0.0021	0.0115	0.0222	-0.0022	0.0114	0.0210	-0.0016	0.0116	0.0226	-0.0017	0.0112	0.0222	-0.0018	0.0120	0.0236
P. radiata (under bark)																
< 10 cm	12	-0.0008	0.0026	0.0031	0.0036	0.0036	0.0039	0.0014	0.0017	0.0020	0.0007	0.0009	0.0010	-0.0004	0.0008	0.0010
10 - 15 cm	16	-0.0006	0.0038	0.0050	0.0045	0.0046	0.0055	-0.0007	0.0030	0.0039	-0.0001	0.0025	0.0035	-0.0010	0.0030	0.0037
15 - 20 cm	22	-0.0013	0.0095	0.0121	0.0039	0.0080	0.0105	-0.0020	0.0074	0.0089	-0.0020	0.0086	0.0105	-0.0018	0.0085	0.0104
20 - 25 cm	17	0.0087	0.0154	0.0237	0.0088	0.0145	0.0225	-0.0012	0.0149	0.0186	0.0004	0.0140	0.0181	0.0070	0.0150	0.0225
25 - 30 cm	18	0.0182	0.0308	0.0429	0.0208	0.0316	0.0442	0.0100	0.0274	0.0385	0.0059	0.0300	0.0379	0.0200	0.0299	0.0426
30 - 35 cm	21	0.0138	0.0480	0.0620	0.0188	0.0513	0.0655	0.0135	0.0497	0.0629	0.0064	0.0453	0.0602	0.0191	0.0511	0.0640
35 - 40 cm	17	-0.0089	0.0450	0.0639	0.0041	0.0422	0.0648	0.0042	0.0413	0.0627	-0.0092	0.0457	0.0642	0.0012	0.0427	0.0630
40 - 45 cm	9	-0.0319	0.0823	0.1057	-0.0056	0.0898	0.1167	-0.0020	0.0917	0.1126	-0.0164	0.0809	0.1010	-0.0204	0.0891	0.1113
45 - 50 cm	17	-0.0060	0.1179	0.1405	0.0080	0.1084	0.1315	0.0209	0.1086	0.1342	0.0198	0.1135	0.1383	0.0058	0.1171	0.1402
50 - 55 cm	10	-0.0018	0.1595	0.2164	0.0003	0.1529	0.1976	0.0099	0.1548	0.2058	0.0352	0.1570	0.2116	-0.0017	0.1641	0.2189
55 -60 cm	4	0.1210	0.2057	0.2214	0.0678	0.1914	0.2071	0.1155	0.1906	0.2082	0.1159	0.2250	0.2420	0.1726	0.1812	0.2198
> 60 cm	5	0.1174	0.1801	0.2003	0.0212	0.1213	0.1370	0.0021	0.1376	0.1544	0.1232	0.1651	0.1861	0.0717	0.1644	0.1917
Total/average	168	0.0067	0.0530	0.0943	0.0095	0.0502	0.0872	0.0077	0.0498	0.0886	0.0096	0.0514	0.0925	0.0097	0.0525	0.0945
P. radiata (over bark)																
< 10 cm	9	-0.0034	0.0053	0.0064	0.0046	0.0046	0.0048	0.0004	0.0012	0.0014	0.0002	0.0006	0.0007	-0.0015	0.0016	0.0019
10 - 15 cm	12	0.0005	0.0058	0.0069	0.0068	0.0068	0.0075	-0.0045	0.0047	0.0064	0.0007	0.0029	0.0040	-0.0012	0.0032	0.0037
15 - 20 cm	10	0.0068	0.0121	0.0162	0.0112	0.0125	0.0158	-0.0096	0.0111	0.0128	0.0036	0.0097	0.0126	0.0032	0.0094	0.0122
20 - 25 cm	10	0.0109	0.0117	0.0207	0.0107	0.0114	0.0181	-0.0249	0.0264	0.0309	0.0034	0.0076	0.0138	0.0061	0.0108	0.0192
25 - 30 cm	9	0.0161	0.0319	0.0375	0.0176	0.0267	0.0333	-0.0368	0.0426	0.0566	0.0061	0.0283	0.0334	0.0225	0.0269	0.0331
30 - 35 cm	17	-0.0081	0.0455	0.0545	-0.0055	0.0423	0.0522	-0.0539	0.0629	0.0839	-0.0155	0.0419	0.0557	0.0018	0.0443	0.0521
35 - 40 cm	9	-0.0164	0.0382	0.0459	-0.0093	0.0394	0.0478	-0.1044	0.1157	0.1281	-0.0158	0.0319	0.0409	0.0011	0.0381	0.0502
40 - 45 cm	7	-0.0666	0.0822	0.1019	-0.0549	0.0742	0.0941	-0.1639	0.1651	0.1969	-0.0587	0.0784	0.0986	-0.0516	0.0822	0.0996
45 - 50 cm	16	-0.0166	0.1074	0.1254	-0.0002	0.1005	0.1158	-0.1607	0.1805	0.2236	0.0033	0.1083	0.1256	0.0014	0.1020	0.1190
50 - 55 cm	10	-0.0108	0.1679	0.2129	-0.0020	0.1616	0.2042	-0.2245	0.2278	0.3260	0.0166	0.1665	0.2139	-0.0114	0.1650	0.2111
55 -60 cm	4	0.0754	0.2021	0.2293	0.0744	0.1892	0.2193	-0.0396	0.0897	0.1189	0.1019	0.2134	0.2420	0.1579	0.1579	0.2041
> 60 cm	5	0.1169	0.1925	0.2191	0.0604	0.1530	0.1821	-0.3187	0.3187	0.3736	0.1261	0.1966	0.2235	0.0579	0.1737	0.2185
Total/average	118	-0.0002	0.0633	0.1054	0.0037	0.0586	0.0978	-0.0750	0.0816	0.1465	0.0041	0.0611	0.1063	0.0052	0.0590	0.1023

In the case of *E. grandis*, which was analysed only for over bark diameters, Equation 2.14a was clearly superior in the final rank. This taper estimation system was superior to the other four at the overall level and for diameter and merchantable height at the *detailed* level (i.e. by percent height for diameter and by *dbh* classes for merchantable height). For total tree volume, although Equation 2.14a showed very small values of B, MAD and SEE (Table 3.6), it was slightly outperformed by Equation 1.16. For this species, an implicit tree volume equation was required for some of the analyses described in Part II. Therefore, Equation 2.11a, which can be rearranged as a tree volume equation, was also retained, despite its relatively poorer performance.

For *P. radiata*, the best two taper estimation systems were Equations 1.16, 2.17b (for under-bark data) and 2.17c (for over-bark data). Equation 1.16 performed better (i.e. lower final rank) than Equation 2.17b for the under-bark subset, whereas Equation 2.17c outperformed Equation 1.16 for the over-bark subset. For both subsets (i.e. over- and under-bark), Equation 1.16 had lower ranks at the overall level and for merchantable height by *dbh* classes (Table 3.7). On the other hand, Equations 2.17b and 2.17c were superior to Equation 1.16 for diameters by percent height. In the case of total tree volume, Equation 1.16 was better ranked than Equation 2.17b but worse ranked than Equation 2.17c. Considering that under-bark estimations are usually considered more relevant and that Equation 1.16 was developed to minimise the level of multicollinearity between its independent variables (Kozak, 1997), Equation 1.16 was selected as the best taper estimation system for *P. radiata*.

Selected equations were also evaluated by visual inspection of plots of residuals against predicted values, *dbh* and total tree height. In all cases, residuals were uniformly distributed around the zero reference line and exhibited no trends with predicted values or explanatory variables. The ranking procedure described above was consistent with the visual evaluation of residual plots. That is, the best-ranked models also presented residual plots in which residuals were more tightly aligned with the zero reference line, exhibiting less variability and no non-normal trends (residuals from some of the worst ranked models did exhibit non-normal trends).

Table 3.7 Sum of ranks and ranks based on the sums (brackets) of selected models

	E. grandis (over bark)					*P. radiata* (under bark)					*P. radiata* (over bark)				
Description	1.8a	2.11a	2.14a	1.16	2.17a	1.8a	2.11a	2.14a	1.16	2.17a	1.8a	2.11a	2.14a	1.16	2.17a
Overall Bias for diameter, V and h90	10 (4)	15 (5)	3 (1)	8 (2)	9 (3)	4 (1)	13 (5)	7 (2)	10 (3)	11 (4)	3 (1)	11 (3)	11 (3)	8 (2)	12 (4)
Overall MAD for diameter, V and h90	10 (4)	12 (5)	6 (1)	8 (2)	9 (3)	12 (2)	12 (2)	7 (1)	7 (1)	7 (1)	12 (5)	11 (4)	9 (3)	7 (2)	6 (1)
Overall SEE for diameter, V and h90	10 (2)	7 (1)	7 (1)	11 (3)	10 (2)	11 (3)	11 (3)	8 (2)	7 (1)	8 (2)	11 (4)	11 (4)	9 (3)	8 (2)	6 (1)
Bias for diameter from ground to top	45 (5)	33 (3)	23 (1)	40 (4)	24 (2)	26 (1)	35 (3)	38 (4)	38 (4)	28 (2)	34 (4)	31 (2)	28 (1)	40 (5)	32 (3)
MAD for diameter from ground to top	34 (3)	40 (5)	23 (1)	39 (4)	29 (2)	38 (3)	44 (4)	25 (1)	37 (4)	23 (1)	45 (3)	49 (4)	19 (1)	33 (2)	19 (1)
SEE for diameter from ground to top	39 (3)	39 (3)	16 (1)	40 (4)	31 (2)	37 (4)	46 (5)	26 (2)	34 (3)	22 (1)	44 (4)	46 (5)	20 (2)	37 (3)	18 (1)
Bias for h90 by *dbh* classes	24 (4)	30 (5)	14 (1)	21 (3)	16 (2)	36 (3)	50 (5)	41 (4)	23 (1)	30 (2)	43 (4)	46 (5)	24 (1)	25 (2)	42 (3)
MAD for h90 by *dbh* classes	23 (3)	30 (4)	15 (1)	22 (2)	15 (1)	37 (2)	48 (4)	44 (3)	14 (1)	37 (2)	47 (5)	46 (4)	28 (2)	19 (1)	40 (3)
SEE for h90 by *dbh* classes	24 (4)	31 (5)	14 (1)	21 (3)	15 (2)	39 (3)	52 (5)	41 (4)	15 (1)	33 (2)	47 (4)	47 (4)	29 (2)	17 (1)	40 (3)
Bias for V by *dbh* classes	17 (1)	20 (2)	22 (4)	21 (3)	25 (5)	39 (3)	39 (3)	32 (1)	35 (2)	35 (2)	37 (4)	33 (3)	51 (5)	32 (2)	27 (1)
MAD for V by *dbh* classes	22 (3)	21 (2)	21 (2)	15 (1)	22 (3)	49 (5)	35 (3)	27 (1)	33 (2)	36 (4)	46 (4)	31 (3)	49 (5)	28 (2)	26 (1)
SEE for V by *dbh* classes	23 (4)	22 (3)	21 (2)	14 (1)	25 (5)	47 (4)	40 (3)	28 (1)	28 (1)	37 (2)	45 (4)	28 (2)	49 (5)	33 (3)	25 (1)
Total	281 (40)	300 (43)	185 (17)	260 (32)	230 (32)	375 (34)	425 (45)	322 (26)	281 (23)	307 (24)	414 (47)	390 (43)	326 (33)	287 (27)	293 (23)
Final Rank	**4**	**5**	**1**	**3**	**2**	**4**	**5**	**3**	**1**	**2**	**5**	**4**	**3**	**1 (2)**	**2 (1)**

3.3 AUTOCORRELATION EFFECTS

Because the taper data are autocorrelated, the consequences from a statistical point of view, which were discussed in Section 1.3, are that standard errors of the estimated coefficients are underestimated and therefore significance tests are unreliable. In order to check the parameter significance in the selected models Proc Mixed (SAS Institute Inc. 1996) was used. This procedure takes account of various measurements taken from each sample. As this procedure is relatively new and its use has seldom been reported in published taper studies (an exception being Tassisa and Burkhart 1998), its essential programming code is exemplified below:

```
proc mixed;
    model y=x1 x2 x3 x4 x5 x6 x7 ;
    repeated / subject=ID type=AR(1);
```

where y is the dependent variable (e.g. diameter); x1 to x7 are the independent variables and ID is a variable identifying each tree. The type=AR(1) option specifies a first-order autoregressive covariance structure, which is the more appropriate for this application (Woollons, R. personal communication).

As the MIXED procedure was conceived to handle linear models (although some complex macros can be used to handle non-linear models as well), the models were transformed into linear models using logarithms. The following linear models correspond to the selected models:

E. grandis

$$\ln(d) = \ln(a_0) + a_1\ln(dbh) + b_1\frac{\ln(X)}{z} + b_2\ln(X)\sqrt{z} + b_3\ln((X)\exp(z) + b_4\ln(X)\left(\frac{dbh}{H}\right)^2 + b_5\ln(X)dbh + b_6\frac{\ln(X)}{H} \quad (3.14a)$$

P. radiata

$$\ln(d) = \ln(a_0) + a_1\ln(dbh) + b_1\frac{\ln(X)}{z} + b_2\ln(X)\sqrt{z} + b_3\ln((X)\exp(z)$$

$$+b_4\ln(X)\left(\frac{dbh}{H}\right)^2+b_5\ln(X)dbh+b_6\frac{\ln(X)}{H} \qquad (3.16)$$

The terms $\ln(a_0)$ in Equations 3.14a and 3.16 correspond to the intercepts of the mixed linear models. Other terms were referred to as variables X1 to X7. When Equation 3.14a was analysed through Proc Mixed, the following output was obtained:

Solution for Fixed Effects						
Effect	Estimate	Std Error	DF	t	Pr > |t|	(Variable)
INTERC.	-0.00996100	0.03267086	278	-0.30	0.7607	$\ln(a_0)$
X1	0.93542944	0.01089487	278	85.86	0.0001	ln(dbh)
X2	0.00126665	0.00010977	5358	11.54	0.0001	ln(X)/z
X3	-0.53388533	0.05029126	5358	-10.62	0.0001	ln(X)*sqrt(z)
X4	0.43519288	0.01872058	5358	23.25	0.0001	ln(X)*/exp(z)
X5	0.16361531	0.01091457	5358	14.99	0.0001	$\ln(X)*(dbh/H)^2$
X6	-0.00337493	0.00048695	5358	-6.93	0.0001	ln(X)*dbh
X7	0.98681560	0.16277036	5358	6.06	0.0001	ln(X)/H

The intercept, which corresponds to the logarithm of the parameter a_0 of Equation 2.14a, was not significant. When the MIXED procedure was asked to solve the parameters of the fixed effects without intercept (by the NOINT option after the model statement) all parameter estimates were highly significant (Pr>|t|, 0.0001). Therefore, the parameter a_0 was removed from Equation 2.14a and the remaining parameters of this equation (referred to as Equation 2.14a* after that modification) were re-estimated through non-linear regression.

$$d=dbh^{a1}\left(\frac{1-\sqrt{z}}{1-\sqrt{p}}\right)^{b_1\sqrt{z}+b_2\exp(z)+b_3/z+b_4(dbh/H)^2+b_5dbh+b_6/H} \qquad (2.14a^*)$$

Statistics of fit after removing the parameter a_0 in Equation 2.14 remained almost unchanged (Table 3.8). Parameter estimates and asymptotic standard errors (ASE) of Equation 2.14a* are shown in Table 3.9.

Table 3.8 Fit statistics of Equations 2.14a and 2.14a* (*E. grandis*).

Equation	Parms.	RSS	MSE	FI	Bias (cm)	MAD (cm)	SEE
2.14a	9	2728.4	0.484	0.993	-0.002	0.481	0.696
2.14a*	8	2728.9	0.484	0.993	-0.003	0.481	0.696

Notes RSS: error sum of square; MSE: mean square error; FI: fit index or percentage of variance explained by the model; other statistics as previously defined.

Table 3.9 Parameter estimates and asymptotic standard errors (ASE) of Equation 2.14a* (*E. grandis*).

	Equation 2.14a*	
Parameter	Estimate	ASE
p	0.2598657	0.00378
a_1	0.9363655	0.00095
b_1	0.0012211	0.00005
b_2	-0.4885558	0.02339
b_3	0.4041347	0.01041
b_4	0.0588137	0.00502
b_5	-0.0007404	0.00020
b_6	2.2971655	0.10171

The segmented taper equation (Equation 2.11a) could not be analysed with Proc Mixed because this procedure cannot handle conditional statements (some parameters of this model apply or do not apply depending on the part of the stem considered). In order to check the significance of the parameters the model was fitted using a single observation per tree randomly selected, which eliminated the autocorrelation problem. All parameters were found significant according to the approximate 95% confidence limits of the parameters (which did not change sign).

When the autocorrelated nature of the data was considered for the selected model for *P. radiata* (Equation 1.16), all parameters were still highly significant (Pr>|t|, 0.0001) for both under- and over-bark data. Therefore, Equation 16 was confirmed as the best model for *P. radiata*.

3.4 TESTING FOR DIFFERENCES BETWEEN FORESTS AND ZONES

So far the models have been compared and evaluated for each subset (i.e. *E. grandis*, over-bark data; *P. radiata* over-bark and under-bark data) overall. However, *E. grandis* data originated in three zones of Uruguay and *P. radiata* data in different forests, which may have induced differences in stem taper. Therefore, differences between forests and zones needed to be tested. This was simply done through F-tests on variance ratios. A *maximum* model, viewed as the sum of all sub-models fitted individually for each forest/zone, was fitted to each subset. The mean square error (MSE) of the maximum model was obtained by dividing the sum of the residual sum of squares (RSS) of all sub-models over the sum of the degrees of freedom of all sub-models. Then, an F-value was calculated by dividing the MSE of the overall model over the MSE of the maximum model. This calculated F-value was compared to tabulated F-values with appropriate degrees of freedom. Due to the large number of degrees of freedom in both models, the tabulated F-values for all typical alpha values corresponded to one and, therefore, the null hypothesis that the models were equal was rejected. Residual variances (MSE) of maximum models were 16.2 and 11% lower than the MSE of overall models for *P. radiata* (Equation 1.16, under-bark subset) and *E. grandis* (Equation 2.14a*) respectively. The reduction of the residual variance of model 2.11a (*E. grandis*) was only 3.1% and, therefore, no attempts were made to localise this model.

In order to cater for the differences between forests/zones, binary indicator (dummy) variables for each zone/region were added to the selected models. The following dummy variables were used:

For each species, the zone/forest with larger number of sampled trees (Zone 7 and Woodhill) were set as default. Each parameter (p) of the selected taper models was substituted with the following expressions:

Dummy variable	D_i=1		D_i=0
	P. radiata	*E. grandis*	
D1	if forest=Dargaville	if Zone=8	
D2	if forest=Whangarei	if Zone=9	
D3	if forest=Maharangi		
D4	if forest=Athenree		otherwise
D5	if forest=Maramuru		
D6	if forest=Riverhead		
D7	if forest=Tairua		

Note: For *P. radiata*, D1, D2 and D3 apply only for the under-bark subset

For *E. grandis* (dob) $p = p_0 + p_1D_1 + p_2D_2$

For *P. radiata* (dub) $p = p_0 + p_1D_1 + p_2D_2 + p_3D_3 + p_4D_4 + p_5D_5 + p_6D_6 + p_7D_7$

For *P. radiata* (dob) $p = p_0 + p_4D_4 + p_5D_5 + p_6D_6 + p_7D_7$

Selected models were re-fitted and those parameters for which the asymptotic 95% confidence interval included zero were promptly removed. Then, the parameters' significances were checked using Proc Mixed. Fit statistics of models with significant dummy variables and overall models are presented in Table 3.10. In general, all statistics improved after the inclusion of dummy variables (i.e. for all statistics but FI, a decrease implies an improvement). The exceptions were the mean residuals (B) for *E. grandis* and for the over-bark subset of *P. radiata*, which increased slightly. To test the null hypothesis that mean residuals or biases were not different to zero, t-tests were conducted according to Rawlings (1988, p.187) and Huang et al. (2000). A t-value was calculated by dividing each mean residual by its standard error, SEE(B). The SEE(B) was calculated by dividing the standard deviation of the residuals, s(B), by the square root of the number of observations. The largest calculated t-value was -1.679 for Equation 2.14a* with dummy variables, which was not sufficient to reject the null hypothesis at the level of α=0.05. All mean residuals presented in Table 3.10 were not statistically different to zero ($p<0.05$), indicating that the models were overall unbiased.

Table 3.10 Fit statistics of selected taper models

			Overall models				Models with dummy variables			
Species	Data type	Equation	FI	B	MAD	SEE	FI	B	MAD	SEE
E. grandis	over bark	2.14a*	0.993	-0.0028	0.481	0.696	0.993	-0.0142	0.468	0.669
E. grandis	over bark	2.11a	0.991	-0.0230	0.539	0.771	-	-	-	-
P. radiata	under bark	1.16	0.988	-0.0207	0.995	1.455	0.990	-0.0081	0.917	1.336
P. radiata	over bark	1.16	0.990	0.0090	1.020	1.534	0.992	0.0096	0.954	1.419

Note FI: fit index or percentage of variance explained by the model

Parameter estimates and asymptotic standard errors (ASE) of selected models including significant dummy variables are presented in Tables 3.11 and 3.12.

Table 3.11 Parameter estimates and ASE of Equation 2.14a* with significant dummy variables (*E. grandis*).

	Equation 2.14a*	
Parameter	Estimate	ASE
a_{10}	0.93406085	0.000287
a_{12}	0.00149625	0.000485
b_{10}	0.00114327	0.000045
b_{12}	0.00065174	0.000065
b_{20}	-0.36990377	0.021411
b_{21}	-0.28113774	0.020119
b_{30}	0.34302149	0.010019
b_{31}	0.14344712	0.009965
b_{32}	0.02684615	0.003954
b_{40}	0.10293224	0.006204
b_{41}	-0.08527580	0.005961
b_{42}	-0.04248401	0.007171
b_{50}	-0.00071793	0.000204
b_{60}	2.02816285	0.098176

For Equation 2.14a* (*E. grandis*), all parameters except b_5 and b_6 were affected by the inclusion of dummy variables. For this model the inclusion of dummy variables reduced the residual variance (MSE) of the model by 7.7%. This reduction was a major part (70%) of the potential maximum reduction that would be achieved by fitting the model to each zone separately.

Table 3.12 Parameter estimates and ASE of Equation 1.16 with significant dummy variables (*P. radiata*).

	under bark data		over bark data	
Parameter	Estimate	ASE	Estimate	ASE
a_{00}	0.9014085	0.00846	0.9551791	0.00864
a_{01}	-0.1446303	0.02902	n/a	
a_{02}	-0.1339108	0.02107	n/a	
a_{04}	0.1300180	0.02301		
a_{10}	1.0101324	0.00442	0.9892889	0.00460
a_{12}	0.0448122	0.00761	n/a	
a_{14}	-0.0358116	0.00596		
a_{15}			-0.0278524	0.01034
a_{16}	-0.0376671	0.01309		
a_{20}	-0.0108138	0.00444	0.0267226	0.00486
a_{21}	0.0733978	0.01481	n/a	
a_{25}			0.0296904	0.01155
a_{26}	0.0414079	0.01450		
b_{10}	0.6494808	0.01144	0.5682421	0.01150
b_{11}	0.4278645	0.03539	n/a	
b_{12}	0.1668831	0.02136	n/a	
b_{13}	0.1015962	0.02956	n/a	
b_{14}	0.1674794	0.01446	0.1842529	0.01386
b_{15}	0.2391976	0.01526	0.2309977	0.01692
b_{16}	0.1183629	0.02196	0.1785866	0.01788
b_{17}	0.2217140	0.01566	0.2042223	0.01564
b_{20}	0.3997434	0.01072	0.4417181	0.01156
b_{21}	-0.1119574	0.04051	n/a	
b_{23}	0.1809637	0.03574	n/a	
b_{30}	-0.4435531	0.02312	-0.0751751	0.02410
b_{32}	0.2383050	0.02591	n/a	
b_{33}	0.3308520	0.04227	n/a	
b_{40}	-0.2977507	0.04184	-0.1718480	0.04199
b_{41}	-0.9944085	0.17381	n/a	
b_{42}	-0.6204726	0.06346	n/a	
b_{43}	-0.6260131	0.07628	n/a	
b_{44}	-0.3951287	0.05518	-0.4360269	0.05425
b_{45}	-0.6846635	0.05917	-0.6206614	0.06512
b_{46}	-0.3164745	0.07796	-0.4691899	0.06785
b_{47}	-0.6924567	0.07351	-0.6078520	0.07361
b_{50}	0.0056339	0.00030	0.0034999	0.00031
b_{56}	0.0027076	0.00037		

In the case of Equation 1.16 fitted to *P. radiata* data, the parameters b_1 and b_4 were highly sensitive to variation in forest. On the other hand, parameters b_2, b_3 and b_5 were constant for almost all forests. The inclusion of dummy variables reduced the residual variance by 15.8 and 14.4%, for the under-bark and over-bark models respectively. When the model for under-bark data was fitted by forest (maximum model) the residual variance was reduced by 16.2% over the overall model. Therefore, around 98% of the potential maximum reduction was achieved by the incorporation of dummy variables.

3.5 VALIDATION OF SELECTED TAPER MODELS

The goodness of fit and prediction ability of selected models (Equations 2.14a* for *E. grandis* and Equation 1.16 for *P. radiata*) after the inclusion of dummy variables were assessed graphically (Figures 3.1, 3.3 and 3.4). The same plots for Equation 2.11.a without dummy variables (*E. grandis*) are displayed in Figure 3.2. The upper left plots in Figures 3.1 to 3.4 show the residuals from developmental subsets against predicted values. The other three plots apply for the validation subsets. The upper right plots in Figures 3.1 to 3.4 show mean residuals (biases) and standard errors (SEE) for diameters by percentage height. The lower plots show mean residuals and standard errors for merchantable height (left) and total tree volume (right) by *dbh* classes.

The scattergrams of residuals against predicted values for all subsets showed little bias. The range of residuals was larger for *P. radiata*, which was expected given the larger number of trees and range of diameters in this dataset. For this species, 80% of residuals ranged between -1.50 and 1.47 cm (for the under-bark subset) and -1.52 and 1.54 cm (for the over-bark subset). In the case of *E. grandis*, 90% of residuals ranged between -1.08 and 1.03 cm for Equation 2.14a* with dummy variables, and between -1.20 and 1.24 cm for Equation 2.11a.

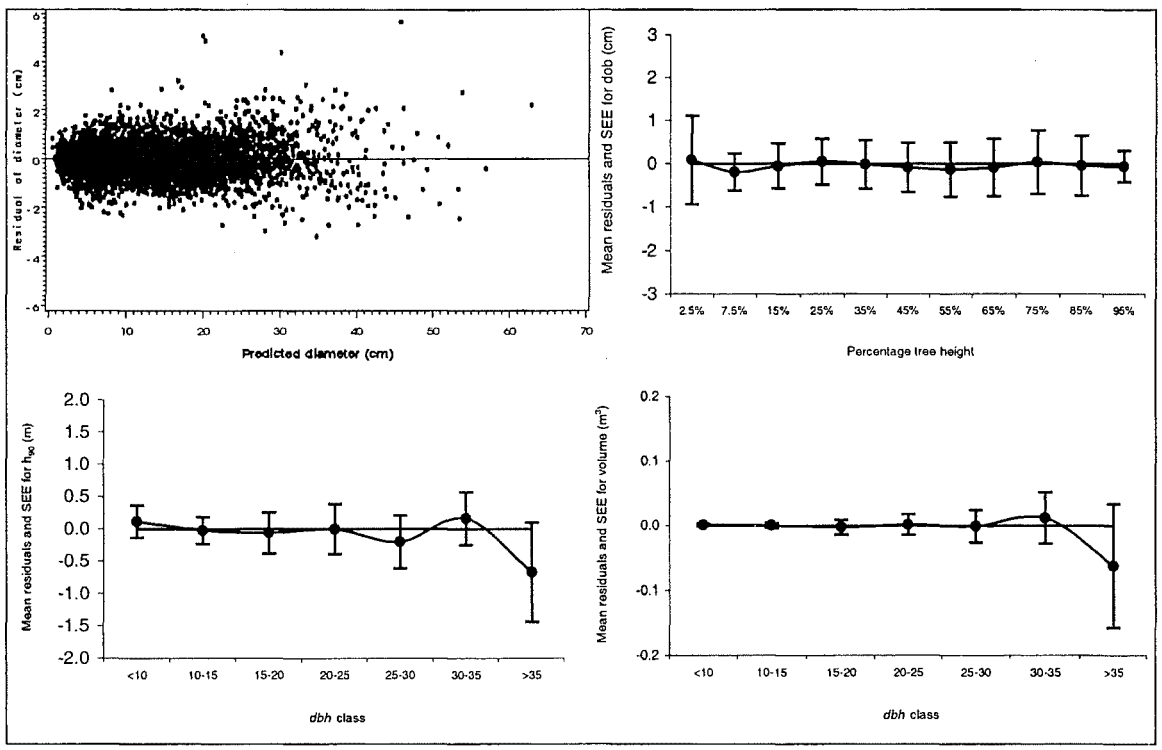


Figure 3.1 Graphical assessment of the goodness of fit and projection ability of Equation 2.14a* with significant dummy variables (*E. grandis*)

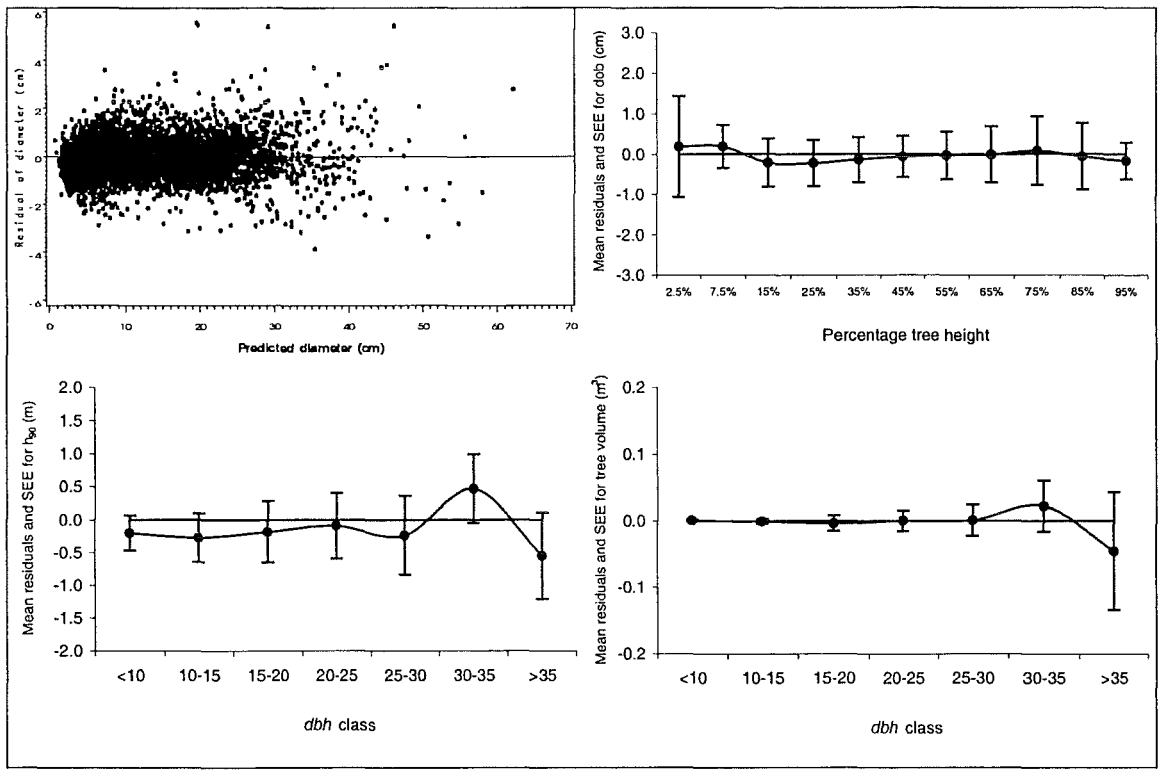


Figure 3.2 Graphical assessment of the goodness of fit and projection ability of Equation 2.11a (*E. grandis*)

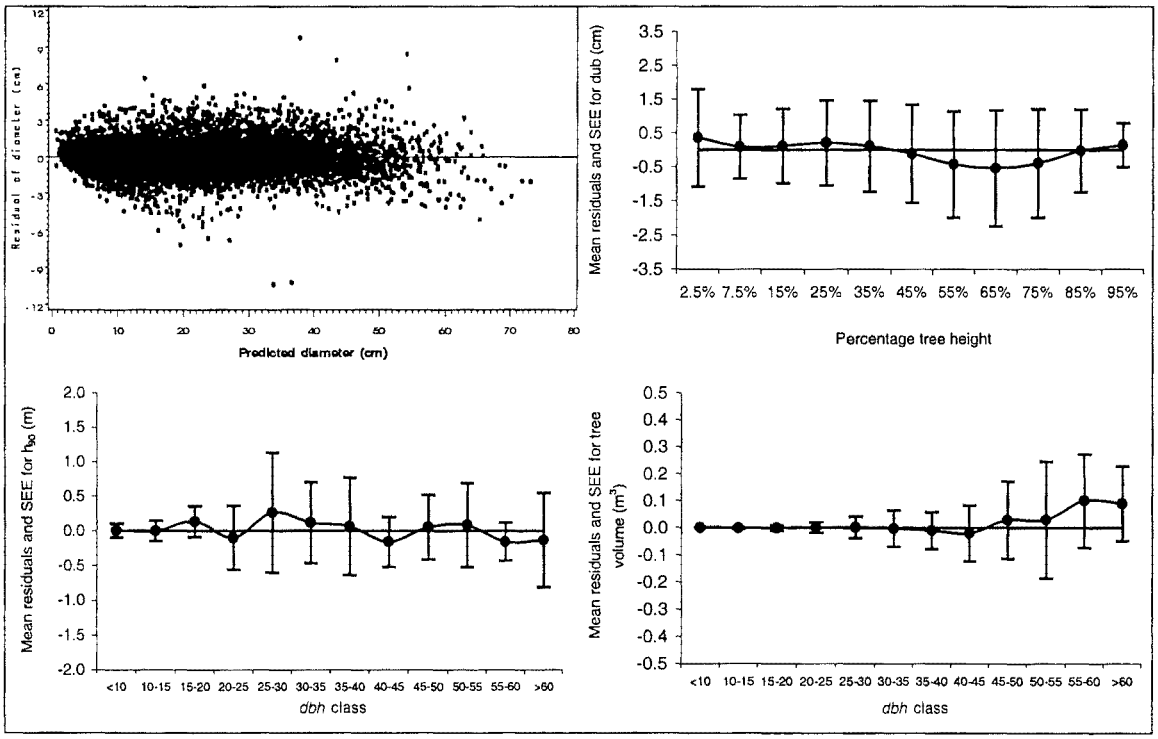


Figure 3.3 Graphical assessment of the goodness of fit and projection ability of Equation 1.16 with significant dummy variables (*P. radiata*, under-bark data)

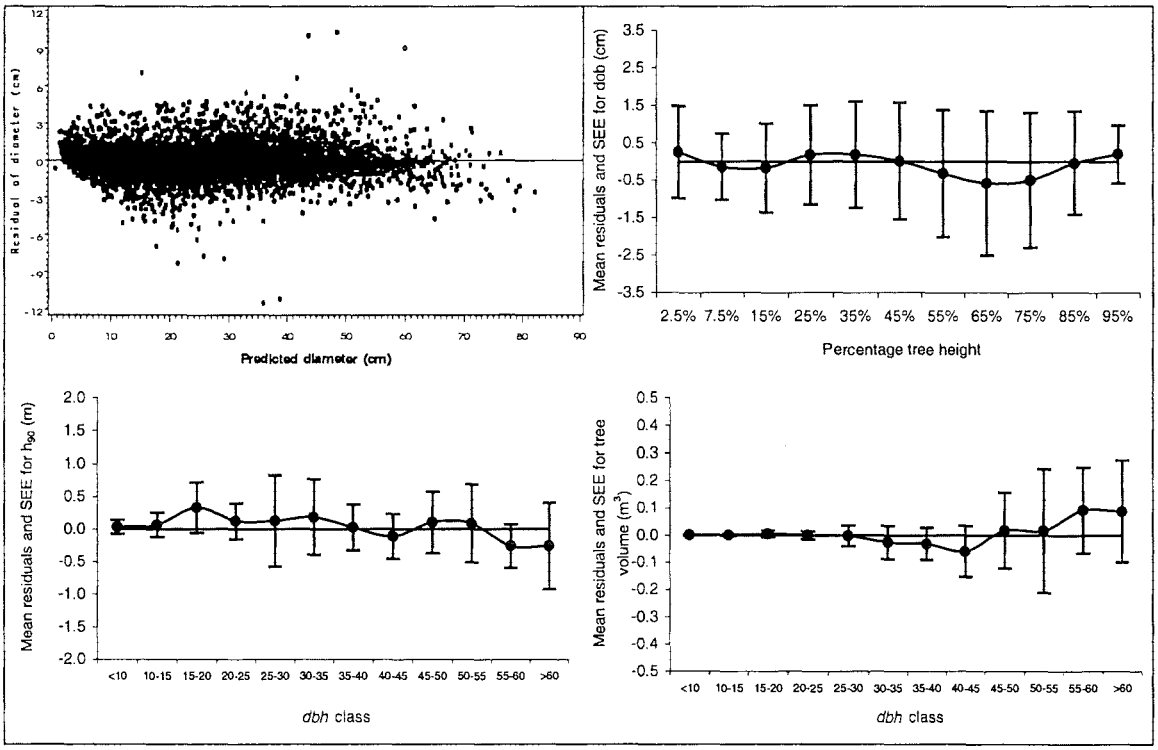


Figure 3.4 Graphical assessment of the goodness of fit and projection ability of Equation 1.16 with significant dummy variables (*P. radiata*, over-bark data)

The plots of mean residuals and standard errors for diameters by percent height (upper right plots in Figures 3.1 to 3.4) indicated that the models exhibited minimal overall bias. However, an undulated pattern around the zero line was evident for *P. radiata*. The same pattern appeared with most of the models tried for this species. Extreme peaks occurred at percent heights between 60-70%, which amounted to -0.54 and -0.58 cm for the under and over-bark subsets respectively (Table 3.14). These negative biases indicated an overestimation of predicted diameters in that part of the stem, which is of less importance as most of the stem's value lies in the lowest third.

Curves of merchantable height biases (lower left plots in Figures 3.1 to 3.4) were close to zero and showed no clear trends, except for Equation 2.11a, which performed poorly (Figure 3.2). Equation 2.14* showed reasonably good performance for merchantable height estimations, although it overestimated this variable at the largest *dbh* class by 0.67 m (less than 2%). This level of inaccuracy is probably within the range of measurement error for heights. As shown in Table 3.13, all other relative biases for height predictions with Equation 2.14a* were lower than 0.85%, with the exception of the bias for the smallest *dbh* class that was 1.33% (0.11 m). The largest relative biases for height predictions in *P. radiata* occurred in the *dbh* class of 15-20 cm., which represented 1.36% (0.14 m) and 3.33% (0.33 m) for the under-bark and over-bark subsets, respectively

Volume estimations were largely unbiased for all *dbh* classes lower than 35-40 cm. However, the magnitude of biases tended to increase with increasing tree size. For *E. grandis*, the largest bias was negative (overestimation) and occurred at the largest *dbh* class (> 35 cm). Although that bias is noticeable in Figures 3.1 and 3.2 (lower right plots) it represented only 2.5% (-0.045 m^3) and 3.4% (-0.062 m^3) of the average tree volume of that *dbh* class, for Equation 2.11a and 2.14a* respectively. In the case of *P. radiata*, biases for the largest *dbh* classes were positive (underestimation). The largest bias for that species occurred in the *dbh* class of 55-60 cm, which was represented by only 3 trees in the validation dataset. Largest volume biases were 0.10 and 0.09 m^3 for under- and over-bark data respectively, which represented 4.4% and 3.5% of average tree volumes for that *dbh* class.

Table 3.13 Biases, percent biases, mean absolute deviations (MAD) and standard error of estimates (SEE) for diameter, merchantable height and total tree volume (*E. grandis*, validation subset).

		Equation 2.14a* (with dummy variables)				Equation 2.11a			
Percent height	n	Bias	Bias (%)	MAD	SEE	Bias	Bias (%)	MAD	SEE
		Diameter estimations (cm)							
0 - 5	184	0.091	0.39	0.719	1.024	0.189	0.81	0.892	1.251
5 - 10	115	-0.191	-0.96	0.327	0.430	0.188	0.94	0.387	0.535
10 - 20	184	-0.054	-0.29	0.393	0.521	-0.218	-1.15	0.448	0.597
20 - 30	189	0.045	0.26	0.392	0.523	-0.218	-1.26	0.445	0.572
30 - 40	188	-0.016	-0.10	0.423	0.561	-0.141	-0.89	0.418	0.561
40 - 50	187	-0.085	-0.59	0.452	0.562	-0.067	-0.47	0.401	0.507
50 - 60	189	-0.131	-1.05	0.473	0.628	-0.030	-0.24	0.446	0.595
60 - 70	182	-0.086	-0.82	0.487	0.662	-0.009	-0.08	0.535	0.694
70 - 80	186	0.039	0.48	0.548	0.734	0.079	0.98	0.658	0.844
80 - 90	181	-0.045	-0.87	0.529	0.689	-0.067	-1.31	0.665	0.825
90 - 100	208	-0.061	-4.25	0.218	0.365	-0.170	-11.75	0.268	0.456
Total/average	1993	-0.040	-0.31	0.452	0.634	-0.052	-0.40	0.506	0.712
dbh class		Merchantable height estimations (m)							
< 10	5	0.114	1.33	0.207	0.251	-0.201	-2.36	0.201	0.266
10 - 15	20	-0.025	-0.20	0.170	0.210	-0.270	-2.21	0.318	0.372
15 - 20	32	-0.063	-0.35	0.264	0.315	-0.184	-1.04	0.404	0.465
20 - 25	25	-0.008	-0.04	0.305	0.386	-0.094	-0.45	0.399	0.501
25 - 30	7	-0.200	-0.81	0.329	0.407	-0.241	-0.98	0.495	0.598
30 - 35	2	0.157	0.50	0.379	0.410	0.471	1.51	0.471	0.521
35 - 40	3	-0.667	-1.97	0.667	0.767	-0.551	-1.63	0.551	0.659
Total/average	94	-0.055	-0.28	0.272	0.348	-0.181	-0.92	0.387	0.469
dbh class		Total tree volume estimations (m^3)							
< 10	5	0.001	4.23	0.002	0.002	0.001	2.91	0.001	0.002
10 - 15	20	0.000	0.27	0.003	0.004	-0.001	-0.70	0.003	0.004
15 - 20	32	-0.002	-1.11	0.009	0.011	-0.004	-1.62	0.010	0.012
20 - 25	25	0.002	0.45	0.013	0.016	0.000	0.03	0.013	0.016
25 - 30	7	-0.001	-0.12	0.020	0.025	0.001	0.12	0.021	0.024
30 - 35	2	0.012	0.96	0.038	0.040	0.022	1.76	0.031	0.038
35 - 40	3	-0.062	-3.39	0.071	0.096	-0.045	-2.50	0.057	0.089
Total/average	94	-0.002	-0.58	0.012	0.022	-0.002	-0.64	0.011	0.021

Table 3.14 Biases, percent biases, mean absolute deviations (MAD) and standard error of estimates (SEE) for diameter, merchantable height and total tree volume (*P. radiata*, Equation 1.16 with dummy variables, validation subsets).

	Under-bark data					Over-bark data				
Percent height	n	Bias	Bias (%)	MAD	SEE	n	Bias	Bias (%)	MAD	SEE
	Diameter estimations (cm)									
0 - 5	372	0.361	0.99	1.136	1.444	314	0.254	0.60	0.918	1.232
5 - 10	242	0.092	0.30	0.748	0.941	184	-0.137	-0.39	0.623	0.883
10 - 20	311	0.115	0.40	0.836	1.103	255	-0.166	-0.52	0.846	1.186
20 - 30	228	0.205	0.77	0.968	1.258	190	0.186	0.63	0.989	1.323
30 - 40	223	0.107	0.46	1.080	1.342	179	0.191	0.73	1.138	1.416
40 - 50	223	-0.114	-0.55	1.096	1.449	171	0.012	0.05	1.178	1.556
50 - 60	203	-0.425	-2.47	1.249	1.571	153	-0.317	-1.63	1.317	1.694
60 - 70	200	-0.542	-3.75	1.306	1.706	149	-0.580	-3.54	1.435	1.926
70 - 80	201	-0.394	-3.70	1.207	1.601	146	-0.493	-4.09	1.307	1.802
80 - 90	193	-0.021	-0.30	0.866	1.223	145	-0.041	-0.52	0.989	1.380
90 - 100	265	0.140	9.17	0.300	0.647	194	0.201	10.78	0.372	0.768
Total/average	2658	-0.001	-0.01	0.968	1.315	2080	-0.043	-0.18	0.978	1.376
dbh class	Merchantable height estimations (m)									
< 10	12	0.000	0.01	0.084	0.099	9	0.033	0.72	0.083	0.108
10 - 15	16	0.001	0.01	0.099	0.144	12	0.069	0.87	0.119	0.187
15 - 20	22	0.135	1.36	0.162	0.223	10	0.331	3.33	0.331	0.391
20 - 25	17	-0.098	-0.65	0.306	0.458	10	0.124	0.83	0.236	0.276
25 - 30	18	0.263	1.35	0.719	0.866	9	0.125	0.64	0.616	0.697
30 - 35	22	0.119	0.55	0.452	0.586	18	0.185	0.85	0.481	0.580
35 - 40	16	0.066	0.27	0.537	0.705	8	0.023	0.10	0.330	0.351
40 - 45	9	-0.160	-0.57	0.282	0.361	7	-0.113	-0.40	0.265	0.342
45 - 50	17	0.053	0.19	0.377	0.466	16	0.103	0.37	0.381	0.469
50 - 55	11	0.082	0.28	0.456	0.607	11	0.089	0.30	0.451	0.599
55 -60	3	-0.158	-0.65	0.259	0.273	3	-0.259	-1.06	0.288	0.336
> 60	5	-0.133	-0.43	0.662	0.680	5	-0.260	-0.83	0.624	0.663
Total/average	168	0.053	0.26	0.362	0.521	118	0.085	0.41	0.352	0.461
dbh class	Total tree volume estimations (m^3)									
< 10	12	0.001	5.23	0.001	0.001	9	0.000	0.73	0.001	0.001
10 - 15	16	-0.001	-1.34	0.002	0.003	12	0.001	0.93	0.003	0.004
15 - 20	22	-0.001	-0.63	0.007	0.010	10	0.004	2.80	0.010	0.012
20 - 25	17	0.000	0.10	0.015	0.020	10	0.000	0.12	0.010	0.014
25 - 30	18	0.002	0.33	0.032	0.040	9	-0.002	-0.41	0.033	0.040
30 - 35	22	-0.003	-0.44	0.051	0.068	18	-0.026	-3.38	0.047	0.064
35 - 40	16	-0.010	-0.92	0.046	0.068	8	-0.031	-2.82	0.051	0.062
40 - 45	9	-0.021	-1.33	0.081	0.103	7	-0.059	-3.46	0.075	0.094
45 - 50	17	0.029	1.50	0.114	0.144	16	0.018	0.81	0.116	0.140
50 - 55	11	0.030	1.18	0.155	0.216	11	0.016	0.56	0.156	0.222
55 -60	3	0.100	4.42	0.165	0.173	3	0.092	3.53	0.152	0.158
> 60	5	0.090	2.38	0.110	0.139	5	0.087	2.00	0.150	0.186
Total/average	168	0.007	0.79	0.049	0.090	118	0.001	0.05	0.060	0.105

CHAPTER 4

BARK MODELS AND THE COMPOSITE APPROACH

4.1 DEVELOPING BARK MODELS

Models discussed in Sections 1.4 and 2.2.3 were fitted to all data having both over-bark and under-bark measurements (133 trees for *E. grandis* and 428 trees for *P. radiata*). Statistics describing the goodness of fit of these models are presented in Table 4.1.

Table 4.1 Fit statistics of bark models

	E. grandis					*P. radiata*				
Equation	parms.	B	MAD	FI	SEE	parms.	B	MAD	FI	SEE
1.22	7	-0.022	0.183	0.999	0.287	7	-0.030	0.375	0.998	0.549
2.23a,b	5	-0.057	0.222	0.998	0.367	4	-0.014	0.362	0.998	0.543
2.24a,b	8	0.006	0.154	0.999	0.274	12	-0.002	0.340	0.999	0.518
2.25a,b	7	0.006	0.154	0.999	0.276	7	-0.003	0.346	0.999	0.525

Notes 1.22: Gordon (1983); 2.23a,b: polynomials in y; 2.24a,b: polynomials in y plus other variables; 2.25a,b: as 2.24a,b but with a limit of 7 parameters; y: (H-h)/H

According to the values of all fit statistics, Equations 1.22 and 2.23a,b showed a poorer performance than Equations 2.24a,b and 2.25a,b. Equation 1.22 was clearly superior to Equation 2.23a (*E. grandis*), although for *P. radiata* Equation 2.23b was slightly superior

to Equation 1.22. On the other hand, polynomials including other relevant variables related to tree size and total tree taper (Equations 2.24a,b and 2.25a,b), showed the best overall performance. When a limit of seven parameters was imposed (Equation 2.25a,b) only marginal changes in fit statistics occurred. Equation 2.25a,b, parameter estimates for which are presented in Table 4.2, was further analysed by visual inspection of residual plots (Figures 4.1 and 4.2).

Table 4.2 Parameter estimates and asymptotic standard errors (ASE) for Equation 2.25a,b

	E. grandis (Equation 2.25a)			*P. radiata* (Equation 2.25b)		
Parameter	variable	estimate	A.S.E.	variable	estimate	A.S.E.
a_0	intercept	0.55906	0.0255	intercept	0.76179	0.0072
a_1	y	-0.22146	0.0454	y^8	-0.16702	0.0029
a_2	ln(H)	0.05126	0.0025	ln(dbh)	-0.01200	0.0021
a_3	1/h	-0.10037	0.0058	ln(H)	0.03894	0.0028
a_4	$1/h^2$	0.02569	0.0022	sqrt(y)	0.08909	0.0054
a_5	$1/h^3$	-0.00211	0.0002	H/sqrt(h)	-0.00094	0.0001
a_6	sqrt(y)	0.43363	0.0643	(dbh/H)/h	0.00784	0.0004

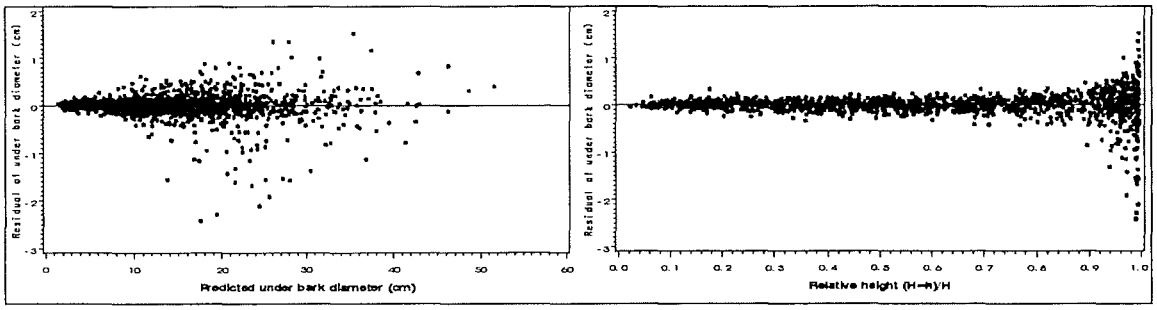


Figure 4.1 Residuals against predicted values (left) and relative height (right) for Equation 2.25a (*E. grandis*)

As expected, the lower part of the tree exhibited greater variation and therefore residuals there were larger. The presence of unequal error variance (heteroscedasticity) in the models led to a consideration of weighted regression analysis. The inverse of *dob* and the inverse of *y* were clear candidates as weighting factors. Both were tried along with a number of variants (such as $1/dob^2$ and $1/y^2$). Although parameter estimates were slightly affected, residual patterns remained virtually unchanged. Variations in residual statistics such as mean residual, standard deviation of residuals, extreme residuals (maximum and

minimum) and mean absolute residual were negligible. Some of them were positively affected by using weighting factors but others (especially the standard deviation of the residuals) were negatively affected. Therefore, it was decided not to use any weighting factor.

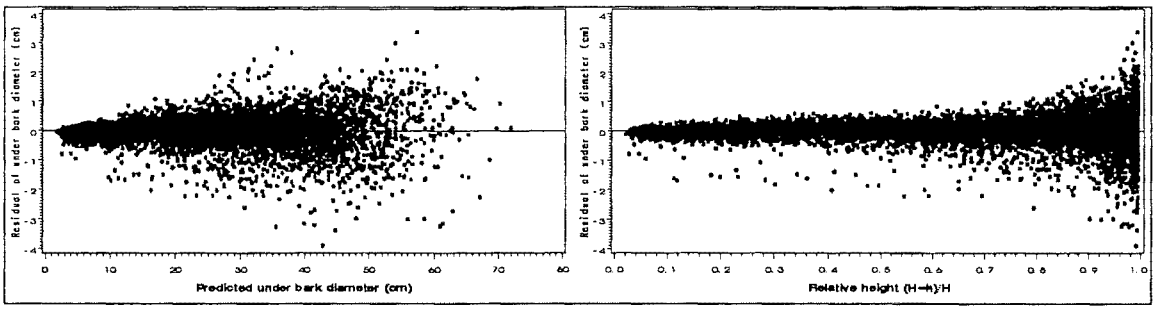


Figure 4.2 Residuals against predicted values (top) and relative height (bottom) for Equation 2.25b (*P. radiata*)

Despite the problem of heteroscedasticity, models exhibited minimal bias. The average error of estimation regardless of sign (MAD) was only 0.15 cm and 0.35 cm for *E. grandis* and *P. radiata* respectively. Ninety percent of residuals ranged between -0.34 to 0.36 cm for *E. grandis* and -0.90 to 0.77 cm for *P. radiata*.

4.2 AUTOCORRELATION EFFECTS

As with the taper models, the effect of the autocorrelated nature of the data was checked for the bark models. Again, this was performed with Proc Mixed with the same details as explained in Section 3.3. Equations 2.25a and 2.25b were fitted as mixed linear models with the dependent variable set as dub^2/dob^2. All parameters were still highly significant (Pr>|t|, 0.0001) for *E. grandis* (Equation 2.25a). In the case of *P. radiata* (Equation 2.25b), although all parameters were significant at the 5% level, the level of significance of the parameter a_2 corresponding to the variable ln(dbh) became 0.0292. This probability of Type I error is much larger than the one set during the process of variable selection ($p<0.0001$). As many other variables were available, a number of combinations of independent variables were tried until all 6 variables and intercept were significant at the level of $p<0.0001$. The best combination of variables when autocorrelation was accounted for during the fitting process was not the same as when it was neglected. The

model with variables selected through Proc Mixed (referred to as Equation 2.25b*) was re-fitted as a non-linear model and the results are shown in Table 4.3.

$$\frac{dub^2}{dob^2}=a_0+a_1y^{10}+a_2\ln(H)+\frac{a_3}{h}+\frac{a_4}{h^2}+a_5\sqrt{y}+a_6\frac{H}{\sqrt{h}} \qquad 2.25b^*$$

Table 4.3 Parameter estimates and fit statistics of Equation 2.25b* (*P. radiata*)

Parameter	variable	estimate	A.S.E.	Fit statistics	
a_0	intercept	0.74553	0.0089		
a_1	y^{10}	-0.19996	0.0046	Mean residual	-0.0043
a_2	ln(H)	0.03497	0.0022	MAD	0.3517
a_3	1/h	0.04259	0.0030	Fit Index	0.9985
a_4	$1/h^2$	-0.00450	0.0004	SEE	0.5321
a_5	sqrt(y)	0.05969	0.0052		
a_6	H/sqrt(h)	-0.00093	0.0001		

Fit statistics of Equation 2.25b* showed minimal changes in relation to those of Equation 2.25b (Table 4.1). Therefore, Equation 2.25b* was selected as the best model for predicting under-bark diameters from estimated over-bark diameters for *P. radiata*, as required for the composite approach. For *E. grandis*, Equation 2.25a with parameters presented in Table 4.2 was confirmed as the best model, after checking the significance of its parameters with Proc Mixed.

4.3 TESTING FOR DIFFERENCES BETWEEN FORESTS AND ZONES

An F-test on the variance ratios between overall models and maximum models (i.e. the sum of all sub-models fitted separately to each zone/forest) was performed to evaluate whether the use of dummy variables was justified. Although the calculated F-values were very low, the large number of degrees of freedom led to reject the null hypothesis that both models were equal. The residual variance of the maximum model of *P. radiata* was just 6.8% lower than the residual variance of the overall model. For *E. grandis*, this reduction was of 12.6%

The same dummy variables defined for the taper models were tried for Equations 2.25a and 2.25b*. Models with dummy variables were fitted by non-linear regression and parameters the 95% confidence limits for which included zero were removed. Then, the parameter significance was further checked using Proc Mixed until all parameters were significant at the 5% level. Fit statistics of models with significant dummy variables and overall models are presented in Table 4.4. All fit statistics improved after the inclusion of dummy variables. Reductions in residual variances (MSE) over overall models were just 5.3 and 3.4% for *E. grandis* and *P. radiata* respectively.

Table 4.4 Fit statistics of the proposed bark models

	Overall models					Models with dummy variables				
Equation	MSE	FI	B	MAD	SEE	MSE	FI	B	MAD	SEE
2.25a	0.076	0.999	0.006	0.154	0.276	0.072	0.999	0.005	0.153	0.268
2.25b*	0.283	0.999	-0.004	0.352	0.532	0.273	0.999	-0.005	0.343	0.523

Student t-tests were conducted to examine whether mean residuals (B) were not different to zero (null hypothesis) for models with and without dummy variables. The null hypothesis could not be rejected in any case indicating that the models were unbiased overall.

Parameter estimates and asymptotic standard errors (ASE) of selected models including significant dummy variables are presented in Table 4.5.

Table 4.5 Parameter estimates and ASE of Equations 2.25a and 2.25b* with significant dummy variables.

	Equation 2.25a (*E. grandis*)		Equation 2.25b (*P. radiata*)	
Parameter	Estimate	ASE	Estimate	ASE
A00	0.5942481	0.02510	0.7773453	0.00911
A01 / A06	-0.1248568	0.01646	-0.1273207	0.01198
A10	-0.2169055	0.04422	-0.1987933	0.00460
A16			-0.0139015	0.00591
A20	0.0415786	0.00274	0.0254200	0.00233
A21 / A26	0.0350038	0.00505	0.0373213	0.00361
A30	-0.1018235	0.00563	0.0415438	0.00295
A40	0.0264980	0.00216	-0.0044541	0.00042
A50	-0.0021998	0.00022	0.0637897	0.00512
A57			-0.0203189	0.00184
A60	0.4284548	0.06266	-0.0009934	0.00008
A66			0.0002216	0.00011
A67			0.0005880	0.00005

4.4 COMPOSITE APPROACH *VS.* MODELLING UNDER-BARK TAPER DIRECTLY

The comparison of the composite approach *versus* the direct modelling of under-bark diameters was conducted only for *P. radiata*. Only 133 trees of *E. grandis* were effectively measured for under-bark diameter, which was deemed insufficient for fitting an appropriate taper function for under-bark diameters directly. On the other hand, 428 trees of *P. radiata* were measured for both under- and over-bark diameters, resulting in an excellent dataset for doing the analyses described below.

The selected bark model (Equation 2.25b*) was used to estimate under-bark diameters (*dub*) from predicted over-bark diameters (*dob*) obtained from the best taper model (i.e. Equation 1.16, with parameter estimates presented in Table 3.12). Two-step estimations of *dub* (composite approach) were compared to actual values and single-step estimations (i.e. obtained from Equation 1.16 fitted to under-bark data).

A number of sub-samples were extracted from the data to simulate and compare different sub-sampling strategies for taking under-bark measurements when a composite approach

is to be adopted. Parameter estimates for the bark model (Equation 2.25b*) were obtained by fitting it to each sub-sample. As some dummy variables were not significant for some sub-samples, the overall model (without dummy variables) was used to evaluate sub-sampling strategies. This was not considered a major drawback, as the inclusion of dummy variables reduced the residual variance of the overall model by only 3.4%. Statistics of residuals for diameters and volumes, computed on 428 trees (7351 observations), are shown in Table 4.6.

Table 4.6 Residual statistics for under-bark diameters and volumes (*P. radiata*)

Residual statistic	direct model+	Composite approach								
		all data	sub-samples of measurements				sub-samples of trees			
			80%	60%	40%	20%	80%	60%	40%	20%
For dub estimations (cm)										
Mean	-0.002	-0.009	-0.010	-0.009	-0.004	-0.009	-0.013	0.003	-0.017	0.002
MAD	1.032	1.010	1.009	1.010	1.009	1.011	1.009	1.010	1.008	1.007
SEE	1.399	1.384	1.384	1.384	1.384	1.388	1.384	1.384	1.383	1.383
Skewness	0.055	-0.187	-0.189	-0.192	-0.185	-0.145	-0.188	-0.194	-0.150	-0.177
Maximum	9.709	10.037	10.050	10.015	10.112	10.238	10.134	10.048	10.175	10.158
Minimum	-9.928	-10.839	-10.846	-10.845	-10.840	-10.772	-10.860	-10.846	-10.772	-10.791
For volume estimations (m^3)										
Mean	0.007	0.005	0.005	0.005	0.006	0.006	0.004	0.005	0.006	0.008
MAD	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.060	0.060
SEE	0.102	0.102	0.102	0.102	0.102	0.102	0.102	0.102	0.102	0.103
Skewness	0.767	0.149	0.147	0.139	0.183	0.280	0.155	0.134	0.276	0.391
Maximum	0.470	0.449	0.448	0.448	0.450	0.454	0.448	0.449	0.453	0.459
Minimum	-0.440	-0.547	-0.547	-0.548	-0.543	-0.533	-0.546	-0.549	-0.533	-0.520
Trees used to fit Equation 2.25b*		428	428	428	428	428	342	257	171	86
Meas. used to fit Equation 2.25b*		7351	5904	4373	2905	1430	5906	4375	2962	1504
Average *dub* meas. per tree		17	14	10	7	3	17	17	17	18

+ The taper model for predicting *dub* directly is Equation 1.16 fitted to 428 trees. Its parameters and significant dummy variables were as follows:

$$a_0 = 0.89086 + 0.11597 * D4$$

$$a_1 = 1.01366 - 0.03247 * D4 - 0.04427 * D6$$

$$a_2 = -0.01081 + 0.0473 * D6$$

$$b_1 = 0.68244 + 0.12715 * D4 + 0.18899 * D5 + 0.15032 * D7$$

$$b_2 = 0.4029$$

$$b_3 = -0.41663$$

$$b_4 = -0.40244 - 0.28135 * D4 - 0.51166 * D5 - 0.3829 * D7$$

$$b_5 = 0.00512 + 0.00325 * D6$$

Only minor differences between the direct modelling of *dub* and the composite approach are evident in the main residual statistics (B, MAD and SEE) presented in Table 4.6. Skewness was the most affected statistic, being increased by the composite approach for diameters and decreased for volumes. Extreme diameter residuals were magnified by the composite approach. Extreme negative volume residuals were also magnified, whereas extreme positive residuals were slightly reduced by the composite approach. No mean residuals in Table 4.6 were statistically different from zero ($p<0.05$) according to the t-tests described by Rawlings (1988) and Huang *et al.* (2000), which were explained in Section 3.4.

Residuals were also assessed by percent height for diameters and by *dbh* class for volumes (Tables 4.7 and 4.8). Biases (mean residuals) from the composite approach with a sub-sample of 40% of trees as compared to modelling *dub* directly are illustrated in Figure 4.3.

Patterns of residual statistics were not clearly reflected by intensity and type of sub-sampling (observations *vs.* entire trees). Nonetheless, the sub-sampling intensity of 20% led to somewhat poorer residual statistics (especially when trees rather than observations were sub-sampled). Residuals from models fitted with under-bark diameter measurements (or bark thickness) from 40% or more of trees or observations were virtually equally to those in the taper model directly fitted to under-bark data (Figure 4.3). The composite approach generated slightly larger diameter biases, but volume estimations from this approach were more accurate than volume estimations obtained from the *direct* approach.

Table 4.7 Biases, MADs and SEEs for under-bark diameters by percent height (*P. radiata*)

Percent height	N	Direct modelling of *dub*			Comp. app. (all data)			Comp. app. (20% meas.)			Comp. app. (40% meas.)			Comp. app. (60% meas.)		
		Bias (cm)	MAD (cm)	SEE (cm)	Bias (cm)	MAD (cm)	SEE (cm)	Bias (cm)	MAD (cm)	SEE (cm)	Bias (cm)	MAD (cm)	SEE (cm)	Bias (cm)	MAD (cm)	SEE (cm)
<5%	1177	0.195	1.148	1.536	0.155	1.012	1.407	0.129	1.007	1.404	0.154	1.009	1.406	0.162	1.012	1.407
5 - 10%	684	-0.175	0.743	0.970	-0.202	0.734	0.966	-0.257	0.734	0.971	-0.192	0.729	0.959	-0.197	0.735	0.966
10 - 20%	895	-0.061	0.839	1.101	-0.195	0.839	1.109	-0.193	0.835	1.106	-0.180	0.836	1.105	-0.195	0.840	1.109
20 - 30%	752	0.158	0.932	1.245	0.143	0.905	1.216	0.174	0.910	1.225	0.157	0.906	1.218	0.141	0.904	1.215
30 - 40%	668	0.169	1.005	1.320	0.294	1.012	1.331	0.322	1.023	1.345	0.304	1.014	1.334	0.292	1.011	1.330
40 - 50%	637	0.033	1.090	1.436	0.175	1.088	1.426	0.195	1.095	1.437	0.181	1.088	1.427	0.172	1.087	1.425
50 - 60%	592	-0.244	1.236	1.711	-0.145	1.213	1.700	-0.132	1.218	1.703	-0.143	1.213	1.699	-0.148	1.213	1.700
60 - 70%	561	-0.394	1.298	1.728	-0.404	1.289	1.754	-0.397	1.293	1.756	-0.404	1.289	1.754	-0.406	1.289	1.755
70 - 80%	548	-0.306	1.221	1.655	-0.380	1.253	1.707	-0.378	1.255	1.708	-0.383	1.253	1.707	-0.382	1.253	1.707
80 - 90%	529	0.051	0.994	1.365	-0.032	1.016	1.399	-0.033	1.016	1.399	-0.035	1.016	1.399	-0.033	1.016	1.399
>90%	308	0.603	0.830	1.049	0.606	0.855	1.080	0.605	0.853	1.079	0.603	0.854	1.079	0.606	0.855	1.080
Total	7351	-0.002	1.032	1.399	-0.009	1.010	1.384	-0.009	1.011	1.388	-0.004	1.009	1.384	-0.009	1.010	1.384

		Comp. app. (80% meas.)			Comp. app. (20% trees)			Comp. app. (40% trees)			Comp. app. (60% trees)			Comp. app. (80% trees)		
<5%	1177	0.163	1.011	1.407	0.161	1.001	1.393	0.093	1.000	1.394	0.194	1.015	1.410	0.196	1.011	1.407
5 - 10%	684	-0.197	0.732	0.963	-0.150	0.705	0.928	-0.248	0.727	0.961	-0.177	0.736	0.965	-0.193	0.729	0.959
10 - 20%	895	-0.197	0.839	1.109	-0.154	0.823	1.094	-0.209	0.836	1.107	-0.191	0.841	1.110	-0.205	0.839	1.109
20 - 30%	752	0.139	0.904	1.215	0.170	0.909	1.226	0.149	0.906	1.220	0.141	0.904	1.214	0.125	0.903	1.213
30 - 40%	668	0.289	1.010	1.330	0.306	1.022	1.343	0.304	1.018	1.339	0.294	1.010	1.329	0.273	1.008	1.327
40 - 50%	637	0.171	1.087	1.425	0.177	1.095	1.436	0.188	1.093	1.434	0.178	1.086	1.424	0.155	1.086	1.424
50 - 60%	592	-0.149	1.213	1.700	-0.150	1.219	1.706	-0.130	1.217	1.703	-0.139	1.211	1.698	-0.165	1.215	1.702
60 - 70%	561	-0.407	1.289	1.755	-0.415	1.297	1.761	-0.387	1.290	1.753	-0.395	1.286	1.751	-0.421	1.293	1.759
70 - 80%	548	-0.383	1.253	1.707	-0.395	1.258	1.712	-0.363	1.252	1.704	-0.371	1.250	1.704	-0.394	1.256	1.711
80 - 90%	529	-0.033	1.016	1.399	-0.046	1.016	1.400	-0.018	1.015	1.398	-0.024	1.016	1.398	-0.042	1.017	1.400
>90%	308	0.605	0.854	1.080	0.597	0.850	1.076	0.616	0.858	1.085	0.612	0.857	1.083	0.601	0.852	1.078
Total	7351	-0.010	1.009	1.384	0.002	1.007	1.383	-0.017	1.008	1.383	0.003	1.010	1.384	-0.013	1.009	1.384

Table 4.8 Biases, MADs and SEEs for under-bark volumes by *dbh* class (*P. radiata*)

		Direct modelling of *dub*			Comp. app. (all data)			Comp. app. (20% meas.)			Comp. app. (40% meas.)			Comp. app. (60% meas.)		
dbh class (cm)	N	Bias (m^3)	MAD (m^3)	SEE (m^3)	Bias (m^3)	MAD (m^3)	SEE (m^3)	Bias (m^3)	MAD (m^3)	SEE (m^3)	Bias (m^3)	MAD (m^3)	SEE (m^3)	Bias (m^3)	MAD (m^3)	SEE (m^3)
<10	36	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001
10 - 15	29	0.002	0.003	0.005	0.001	0.003	0.005	0.001	0.003	0.005	0.001	0.003	0.005	0.001	0.003	0.005
15 - 20	39	0.002	0.009	0.011	0.004	0.009	0.011	0.003	0.009	0.011	0.004	0.009	0.011	0.004	0.009	0.011
20 - 25	24	-0.003	0.013	0.017	0.000	0.014	0.018	-0.001	0.014	0.018	0.000	0.014	0.018	0.000	0.014	0.018
25 - 30	35	-0.008	0.021	0.027	-0.005	0.021	0.028	-0.005	0.021	0.028	-0.004	0.021	0.028	-0.005	0.021	0.028
30 - 35	47	-0.019	0.037	0.050	-0.015	0.037	0.051	-0.015	0.037	0.051	-0.015	0.037	0.051	-0.015	0.037	0.051
35 - 40	48	-0.016	0.057	0.076	-0.013	0.058	0.077	-0.013	0.058	0.077	-0.013	0.058	0.077	-0.013	0.058	0.077
40 - 45	42	0.005	0.074	0.099	0.006	0.076	0.099	0.007	0.076	0.099	0.007	0.076	0.099	0.006	0.076	0.099
45 - 50	52	0.001	0.100	0.123	0.000	0.098	0.124	0.003	0.098	0.124	0.001	0.098	0.124	0.000	0.098	0.124
50 - 55	36	0.023	0.124	0.162	0.021	0.124	0.168	0.026	0.125	0.168	0.023	0.124	0.168	0.021	0.124	0.168
55 - 60	23	0.106	0.160	0.197	0.085	0.153	0.188	0.090	0.156	0.191	0.086	0.154	0.189	0.084	0.153	0.188
> 60	17	0.077	0.190	0.237	0.027	0.191	0.240	0.036	0.191	0.240	0.029	0.191	0.240	0.026	0.191	0.240
Total	428	0.007	0.060	0.102	0.005	0.060	0.102	0.006	0.060	0.102	0.006	0.060	0.102	0.005	0.060	0.102

		Comp. app. (80% meas.)			Comp. app. (20% trees)			Comp. app. (40% trees)			Comp. app. (60% trees)			Comp. app. (80% trees)		
<10	36	0.000	0.001	0.001	-0.001	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001
10 - 15	29	0.001	0.003	0.005	0.000	0.004	0.005	0.001	0.003	0.005	0.001	0.003	0.005	0.001	0.003	0.005
15 - 20	39	0.004	0.009	0.011	0.003	0.009	0.011	0.003	0.009	0.011	0.004	0.009	0.012	0.004	0.009	0.011
20 - 25	24	0.000	0.014	0.018	-0.001	0.014	0.018	-0.001	0.014	0.018	0.001	0.014	0.018	0.000	0.014	0.018
25 - 30	35	-0.005	0.021	0.028	-0.006	0.021	0.028	-0.006	0.021	0.028	-0.004	0.021	0.028	-0.005	0.021	0.028
30 - 35	47	-0.015	0.037	0.051	-0.015	0.037	0.051	-0.016	0.037	0.051	-0.014	0.037	0.051	-0.016	0.037	0.051
35 - 40	48	-0.013	0.058	0.077	-0.012	0.058	0.077	-0.014	0.058	0.077	-0.013	0.058	0.077	-0.014	0.058	0.077
40 - 45	42	0.006	0.076	0.099	0.009	0.075	0.099	0.006	0.076	0.099	0.006	0.076	0.099	0.005	0.076	0.099
45 - 50	52	0.000	0.098	0.124	0.006	0.098	0.124	0.002	0.098	0.124	0.000	0.098	0.124	-0.001	0.098	0.124
50 - 55	36	0.021	0.124	0.168	0.031	0.125	0.169	0.026	0.124	0.168	0.021	0.124	0.168	0.021	0.124	0.168
55 - 60	23	0.084	0.153	0.188	0.096	0.159	0.194	0.089	0.156	0.191	0.085	0.153	0.188	0.084	0.153	0.188
> 60	17	0.026	0.191	0.240	0.044	0.191	0.241	0.035	0.190	0.240	0.026	0.191	0.240	0.026	0.191	0.240
Total	428	0.005	0.060	0.102	0.008	0.060	0.103	0.006	0.060	0.102	0.005	0.060	0.102	0.004	0.060	0.102

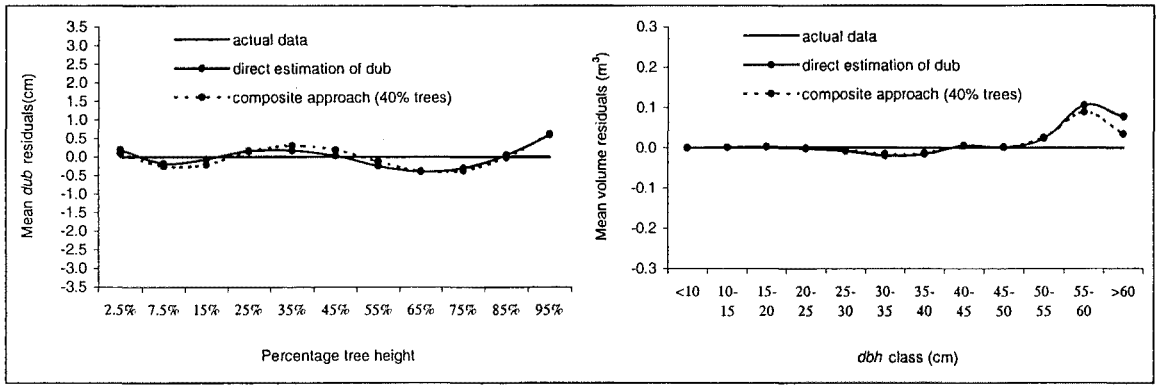


Figure 4.3 Mean residuals of under-bark diameters by percent height (left) and tree volume residuals by *dbh* class (right)

Another way to assess whether the results from the composite approach were substantially different to the results obtained by directly modelling under-bark diameters is through the inspection of residual plots (Figure 4.4). Residual plots were virtually equal with only minor differences in a few extreme observations. The distribution of residual diameters around the zero line was slightly better for the composite approach (left plots in Figure 4.4).

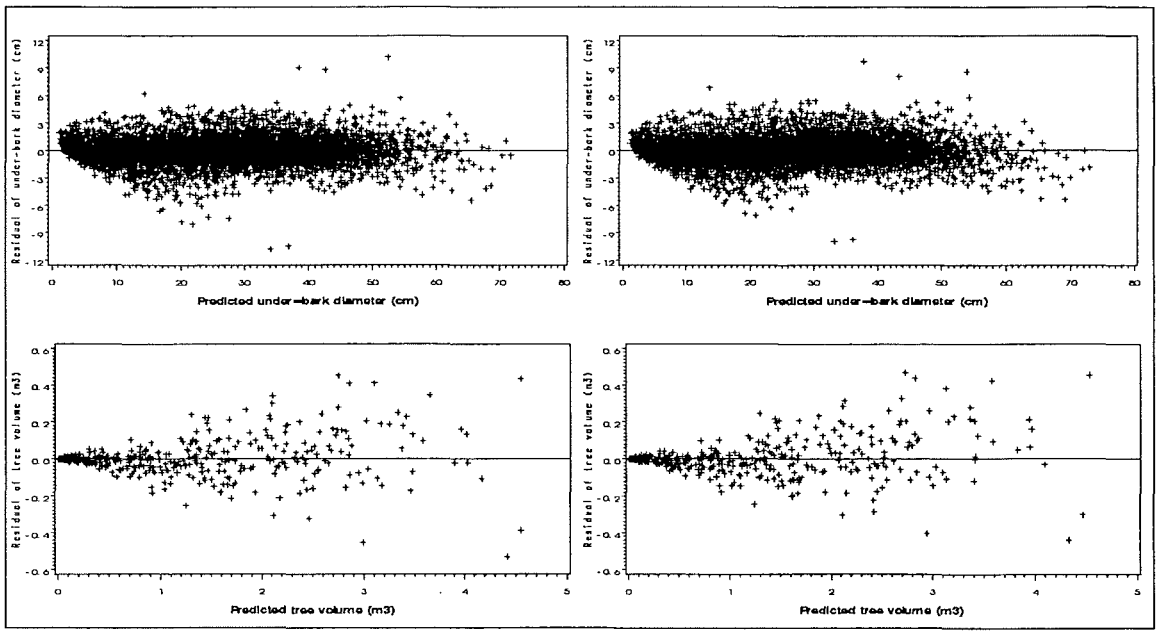


Figure 4.4 Residuals of diameters (top) and tree volumes (bottom) against predicted values obtained from the composite approach with 40% trees (left) and from directly modelling *dub* (right).

CHAPTER 5

DISCUSSION AND CONCLUSIONS

5.1 TAPER MODELS

During the first screening of candidate models only the overall performance for diameter predictions was considered. If this criterion was regarded as definitive, as in many published studies (e.g. Amidon 1984; Candy 1989; Perez *et al.* 1990) the selected models would have been Equations 2.14a, 2.17b and 2.14c for *E. grandis*, under-bark subset of *P. radiata* and over-bark subset of *P. radiata* respectively. Equation 1.16, which was the selected model for *P. radiata,* would have been rejected. This illustrates the relevance of doing a careful scrutiny of the performance of taper models, not only considering diameter predictions, but also volume and height estimations.

Variable-exponent models clearly presented the best performance among the models tested. The selection of the variables for the exponent was successfully done through the MAXR method that selected the best combination of variables for a given number of parameters. Variable-exponent models cannot be inverted to predict heights, nor can they be explicitly integrated to calculate volumes. Therefore, height predictions were obtained with the Newton-Raphson algorithm using Proc Model in the SAS package. Total volumes were obtained by summation of sectional volumes, which were calculated with the conoid formula as *actual* volumes.

The performance of the well-known segmented polynomial by Max and Burkhart (1976) was disappointing, ranking at the 12th, 13th and 17th positions out of 20 models evaluated (Table 3.2). The variant proposed here (Equation 2.11a), with the addition of ln(*dbh*) to the third and fourth terms of Equation 1.11, performed better than the original model and ranked between the 7th and 8th position. This variant reduced to a large extent the undesirable pattern of the residuals with *dbh* (Figure 5.1) and improved all fit statistics.

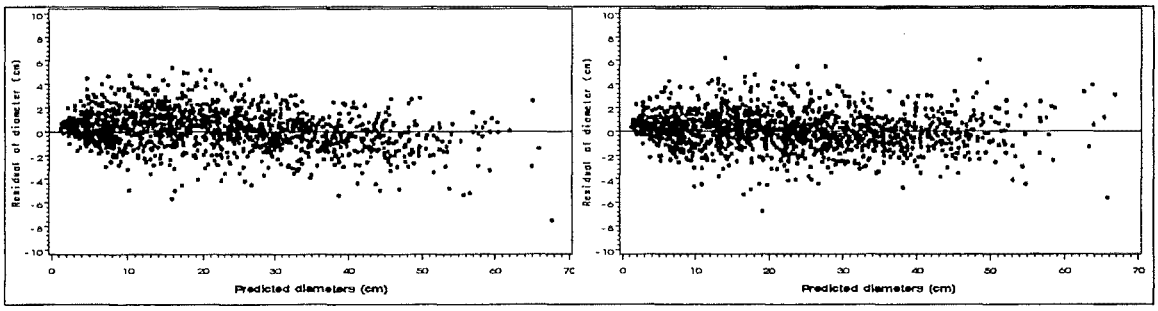


Figure 5.1 Diameter residuals of Equation 1.11 (left) and Equation 2.11a (right). Residuals computed on 20% of the observations in the developmental subset of *P. radiata* for under-bark diameters.

Although the performance of Equation 2.11a was poorer than the performance of variable-exponent models, this model has the advantage that can be explicitly inverted to estimate heights and integrated into a total tree volume equation. As an explicit tree volume equation was required for the second part of this study (Part II), this model was extensively evaluated and proved satisfactory, particularly for volume estimations (Figure 3.2).

Some of the proposed variants for high order polynomials on percent height including *dbh* and H interactions also demonstrated very good performance, especially Equation 1.8a. Variants were proposed to overcome the fact that some models were not appropriate for over-bark diameters. However, these variants performed better than the original formulations, even for under bark diameters. Equation 1.8a has the advantage that its parameters can be obtained by linear regression analysis, although this also applies for variable-exponent models. Its drawbacks are that it cannot be inverted to predict heights, nor can it be explicitly integrated to calculate volumes. These are the same problems as variable-exponent models, the outstanding performance of which would lead to their being preferred to high order polynomials such as Equation 1.8a.

The proposed variable-exponent models proved satisfactory not only for diameter predictions, but also for height and volume estimations. Their predictions were validated against independent datasets (formed by approximately 25% of the data) by assessing their performance in a detailed fashion. Mean residuals, standard errors and mean absolute deviations were computed and scrutinised by percent heights (for diameter predictions) and diameter classes (for merchantable height and total volume estimations). The performance of the selected models was considered not only satisfactory, but also the best attainable given the data available. In order to assure better results when using the models, parameter estimates were refined by using complete datasets, i.e. combining developmental and validation subsets (Tables 5.1 and 5.2).

The models should be used only for the forests or zones contained in the datasets, or very similar ones. Tree sizes represented in the data were very well balanced for *P. radiata*, over a range of small trees (4-cm dbh, 2-m height) to very large trees (70-cm dbh, 40-m height). Therefore, the applicability of Equation 1.16 for *P. radiata* is virtually unrestricted over the complete range of tree sizes, for most common situations. In the case of *E. grandis*, though, the range of tree sizes contained in the dataset was narrower, with trees ranging from 6-cm *dbh* and 6-m height to 50-cm *dbh* and 40-m height. The use of the models proposed for that species (Equations 214a* and 2.11a) should be restricted to trees within the mentioned range, although minor extrapolations are unlikely to be severely biased.

Table 5.1 Parameter estimates of proposed variable-exponent taper models obtained from complete datasets.

	Dummy	*E. grandis* (dob), Eq. 2.14a*		*P. radiata* (dub), Eq. 1.16		*P. radiata* (dob), Eq. 1.16	
Parameter	variable	overall	w/dummy var.	overall	w/dummy var.	overall	w/dummy var.
a_0				0.885193063	0.902120144	0.951744621	0.950628549
a_{01}	D1				-0.140072526		n/a
a_{02}	D2	n/a			-0.129390859		n/a
a_{04}	D4				0.108426195		
a_1		0.934295631	0.933668359	1.016989094	1.010504843	0.993201801	0.991881749
a_{12}	D2		0.002117124		0.043362944		n/a
a_{14}	D4		n/a		-0.030158571		
a_{16}	D6		n/a		-0.038549269		
a_2				-0.012081384	-0.010868974	0.02353432	0.02533074
a_{21}	D1	n/a			0.071139201		n/a
a_{26}	D6				0.042274787		
b_1		0.001267134	0.001201238	0.766271134	0.686454592	0.687169407	0.606116263
b_{11}	D1				0.398839239		n/a
b_{12}	D2		0.000601018		0.134735254		n/a
b_{13}	D3		n/a		0.113362218		n/a
b_{14}	D4		n/a		0.120668999		0.135759235
b_{15}	D5		n/a		0.181967549		0.196582796
b_{16}	D6		n/a		0.024159566		0.094713447
b_{17}	D7		n/a		0.142134231		0.131546475
b_2		-0.454189421	-0.356705152	0.430444112	0.405257835	0.448087921	0.450848852
b_{21}	D1		-0.266888729		-0.119300696		n/a
b_{23}	D3		n/a		0.158667935		n/a
b_3		0.389806027	0.337729415	-0.315983839	-0.417672888	-0.040364417	-0.052763032
b_{31}	D1		0.134283903				
b_{32}	D2		0.021407012		0.224333408		n/a
b_{33}	D3		n/a		0.300319031		n/a
b_4		0.057470051	0.097938888	-0.704284674	-0.450157885	-0.484112836	-0.329585774
b_{41}	D1		-0.080660865		-0.862803599		n/a
b_{42}	D2		-0.037132566		-0.489386536		n/a
b_{43}	D3		n/a		-0.596171968		n/a
b_{44}	D4		n/a		-0.231198397		-0.268205388
b_{45}	D5		n/a		-0.458933013		-0.46669631
b_{46}	D6		n/a				-0.19928509
b_{47}	D7		n/a		-0.324245623		-0.28327461
b_5		-0.00057686	-0.000531735	0.00456235	0.005259536	0.002997722	0.003110734
b_{56}	D6		n/a		0.002398767		
b_6		2.313229194	2.100623417	n/a		n/a	
p		0.267525693	0.267525693	n/a		n/a	

Note: n/a indicates that the parameter is not applicable for the selected model

Table 5.2 Parameter estimates and standard errors of Equation 2.11a for *E. grandis* obtained from complete datasets (under-bark diameters not measured were estimated with Equation 2.25a with parameters in Table 4.5)

	Over bark		Under bark	
Parameter	Estimate	Std. Error	Estimate	Std. Error
b_1	-3.5581	0.0293	-2.7339	0.0241
b_2	1.6909	0.0163	1.2704	0.0137
b_3	-0.4572	0.0061	-0.3684	0.0058
b_4	50.4712	2.0019	40.3051	2.4166
a_1	0.7639	0.0034	0.6865	0.0049
a_2	0.0626	0.0010	0.0530	0.0014

5.2 BARK MODELS AND THE COMPOSITE APPROACH

The composite approach for developing over- and under bark taper equations has been used for eucalypt species in New Zealand (Gordon *et al.* 1995, 1999). With this approach, only a sub-sample of the trees needs to be measured for bark thickness or *dub*, saving resources when collecting taper data. Despite the appeal of this approach, its validity has not yet been tested. A simple way of evaluating the procedure is to compare under-bark predictions (diameters and volumes) against estimations obtained through directly modelling the under-bark profile, as is the usual convention. The dataset of *P. radiata* available for this study, which contained 428 trees with both over-bark and under-bark measurements, was adequate for the proposed analysis.

It was envisaged that the bark model to be used was a crucial factor affecting the overall performance of the method. Polynomials on relative height (y) enriched with other explanatory variables related to tree size and total tree taper (*dbh*/H), were found very useful. Models of that kind, restricted to 7 parameters (Equations 2.25a and 2.25b for *E. grandis* and *P. radiata* respectively) were found satisfactory according to graphical plots of residuals (Figures 4.1 and 4.2) and fit statistics.

The composite approach was found acceptable not only at the overall level (Table 4.6), but also when diameter predictions were scrutinised by percent height (Table 4.7) and

volume estimations were analysed by *dbh* classes (Table 4.8). Furthermore, residuals from the composite approach with 40% of trees sub-sampled for dub measurements showed a better distribution than residuals obtained by modelling under-bark diameters directly (Figure 4.4). Although diameter biases were slightly larger for the composite approach at a few percent heights, volume biases were smaller, especially for the largest *dbh* classes.

Different sub-sampling types and intensities produced similar results, except for the lowest sampling intensity (20%) especially when trees rather than observations were extracted. According to the results in Tables 4.6 to 4.8, measuring under-bark diameters in 40% of trees for fitting a bark model would be sufficient. That 40% represented less than 3000 under bark measurements. When the sub-sampling was performed by trees (i.e. *dub*'s at all measuring points recorded for each sample tree), only 171 trees were considered with 17 measuring points along the stem on average. These results suggest that recording under-bark/over-bark diameter pairs out of approximately 150 trees or 2500 measurements would be sufficient for reliably modelling a relationship between over- and under-bark diameters along the stem.

The bark thickness at the lower part of the stem exhibited large variability as indicated by the broader range of residuals at that section (Figure 4.1). As this region is the most valuable, it seems wise to sample it more intensively. A reasonable scheme for collecting data for taper modelling when a composite approach is to be followed would include measurements of under-bark diameters (or bark thickness) of all trees at breast height (1.30 or 1.40 m) and at a lower point (e.g. 0.65 or 0.70 m), which can easily be done before felling the tree. In a proportion of 20 to 40% of the trees (depending on the total number of trees to be felled) under-bark diameters or bark thickness at all measuring points defined for over-bark measurements (e.g. every one metre) would be recorded.

Measurements of under-bark diameters (or bark thickness) may be subject to more frequent and/or larger errors than over-bark measurements. For instance, the use of bark gauges for measuring bark thickness can introduce bias related to operator experience, operator fatigue, season, and the measurement position of the stem (Gordon 1983). If fewer under-bark measurements were to be taken it would be possible to reduce

measurement errors (either systematic or random). Consequently, the results of the composite approach might be comparatively better.

5.3 AUTOCORRELATED DATA AND REGIONAL VARIATIONS

Statistical inference concerning the significance of parameter estimates is not valid when least-squares methods are applied to autocorrelated data. In this study, the significance of all parameters in the proposed models (both taper and *bark* models) was checked by accounting for the autocorrelated nature of the data. This was conducted through Proc Mixed (SAS, 1996) using *restricted maximum likelihood* (REML) methods implemented with a Newton-Raphson algorithm to solve the parameters. Parameter significance in Equation 2.11a, which could not be analysed with Proc Mixed, was checked with an autocorrelation free dataset formed with a single observation per tree. Accounting for the autocorrelated nature of the data generally used for developing taper models is largely unreported in the literature, and it was a simple and powerful technique that reinforced the validity and stability of the models developed in this study.

The presence of local variations between zones (*E. grandis*) or forests (*P. radiata*) in stem profile was tested and found relatively important. The dummy-variable approach was utilised to allow for locality variations, which allowed tailoring the models to suit each situation in the best possible way. The major reduction in the residual variance was achieved for the under-bark taper model of *P. radiata*, which represented a 15.8% reduction over the overall model. Only minor reductions in residual variances were achieved for bark models.

5.4 DIAGRAMMATIC REPRESENTATION OF THE PROPOSED MODELS

Tree profiles obtained from the proposed taper equations for *P. radiata* (Equation 1.16, parameters in Table 5.1) are graphically displayed in Figure 5.2. The plot on the left in Figure 5.2 shows the under-bark and over-bark tree profiles of two trees with different

size. The areas between the two lines indicate bark thicknesses. The plot on the right in Figure 5.2 shows under-bark profiles for a range of tree sizes spanning the data set.

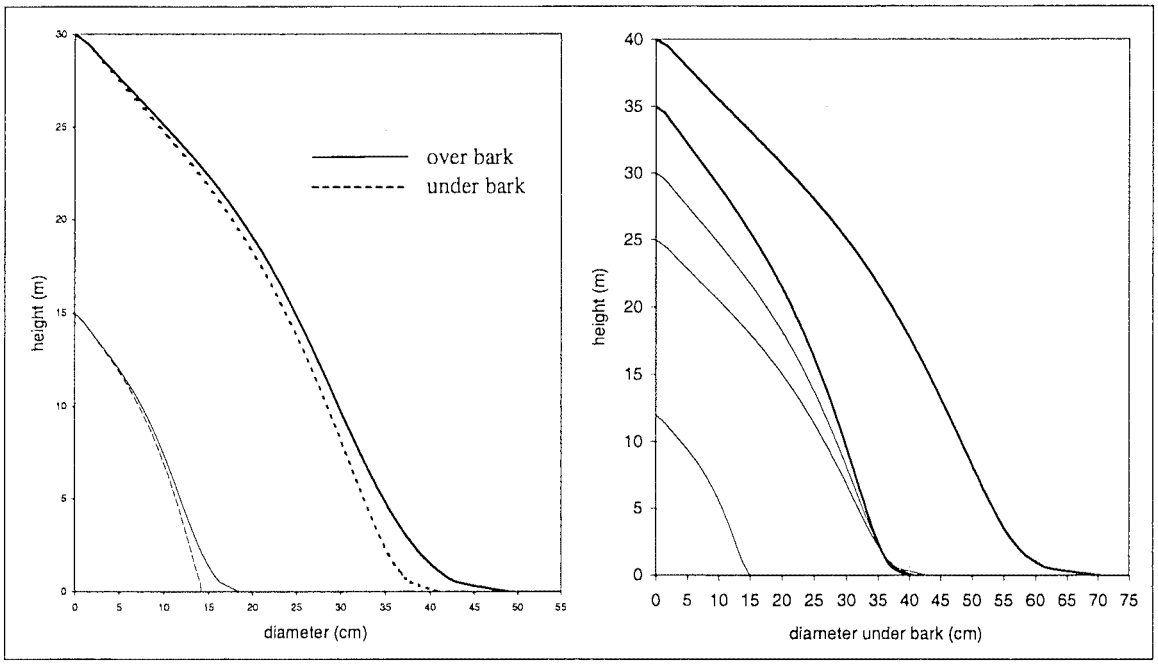


Figure 5.2 Left: under- and over-bark profiles of a small tree (*dbh* 15 cm, H 15 m) and a large tree (*dbh* 40 cm, H 30 m). Right: under-bark profiles of a range of tree sizes (*dbh* from 15 to 65 cm, H from 12 to 40 m). *P. radiata*, Woodhill forest

Analogous plots for *E. grandis* are shown in Figure 5.3. Over bark diameters were obtained from Equation 2.14a* with dummy variables (parameters in Table 5.1). Under-bark diameters were estimated from predicted over-bark diameters with the selected bark equation (Equation 2.25a with dummy variables), which parameters were presented in Table 4.5.

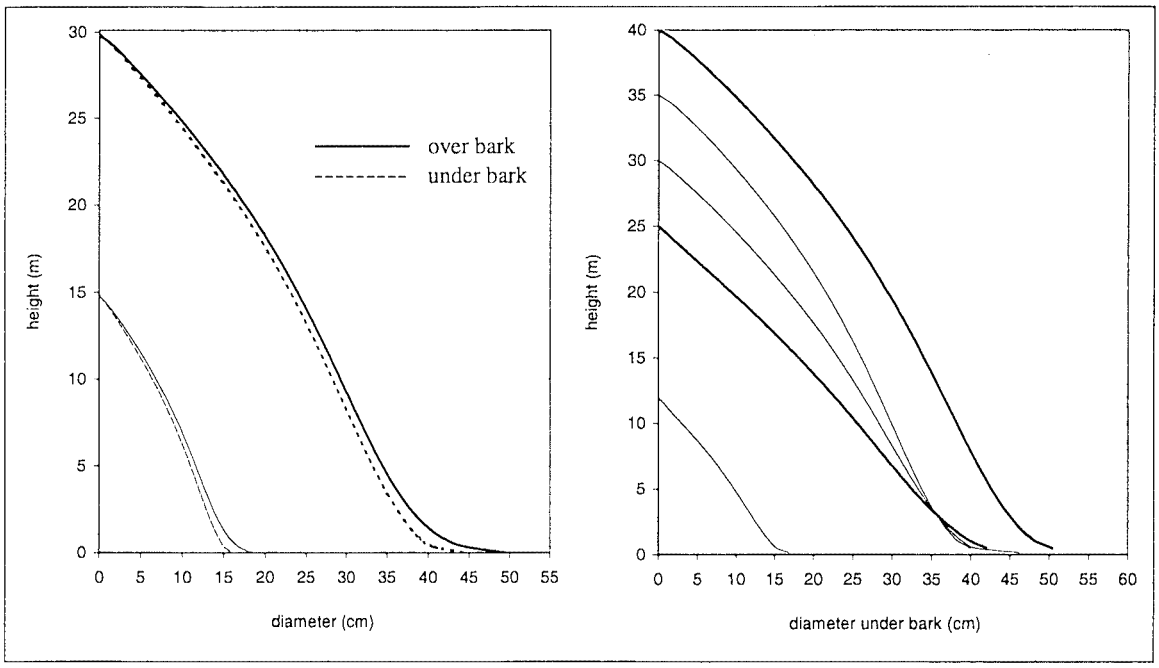


Figure 5.3 Left: under- and over-bark profiles of a small tree (*dbh* 15 cm, H 15 m) and a large tree (*dbh* 40 cm, H 30 m). Right: under-bark profiles of a range of tree sizes (*dbh* from 15 to 50 cm, H from 12 to 35 m). *E. grandis*, Zone 7

Tree profiles displayed in Figures 5.2 and 5.3 illustrate the logical features of the models. The bark thickness increased with tree size and decreased from ground level upwards. The bark of *P. radiata* was thicker than the bark of *E. grandis*, especially at the lower part of the stem. Under bark diameters were always smaller than over-bark diameters at any given height. This was true when both profiles were independently modelled (Figure 5.2, *P. radiata*) and when under-bark diameters were estimated from predicted over-bark diameters (Figure 5.3, *E. grandis*). Predicted diameters at tree top were zero and increased monotonically downwards. Predicted over-bark diameters at breast height (1.30 m for *E. grandis* and 1.40 m for *P. radiata*) coincided with *dbh*. The butt swelling was more pronounced for larger trees.

Under-bark profiles of trees with the same *dbh* but different heights (plots on the right in Figures 5.2 and 5.3) crossed each other above breast height. For trees with 40-cm dbh and heights of 25, 30 and 35 m, under-bark profiles crossed at approximately 3.00 m height for *E. grandis* and 1.70 m height for *P. radiata*. This indicated that the under-bark profiles of the two species differed in the lower part of the stem, which was confirmed by plotting under-bark profiles of a tree of the same size for both species (Figure 5.4)

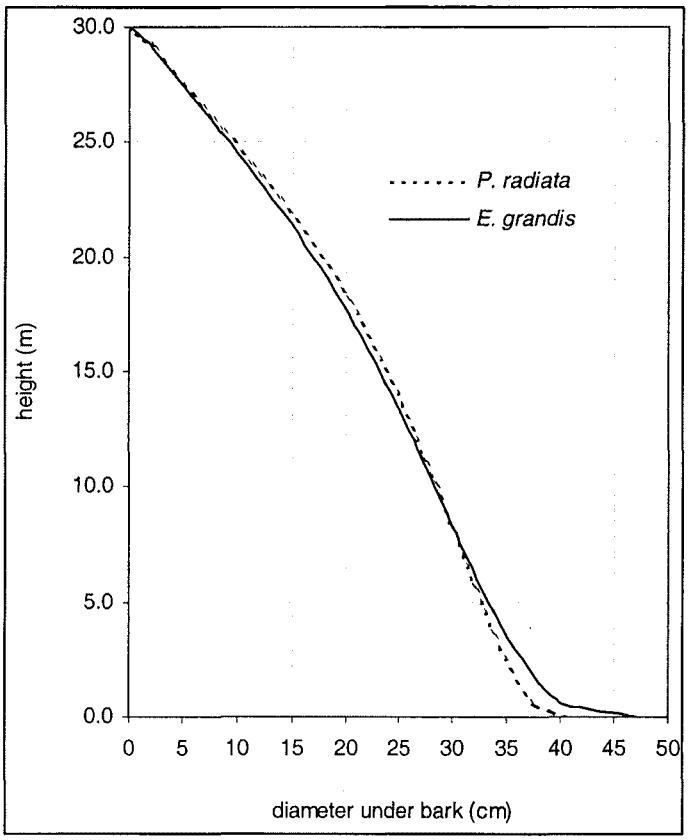


Figure 5.4 Under-bark profiles of a 40-cm dbh, 30-m height tree for both species.

It appeared from Figure 5.4 that *E. grandis* allocated comparatively more volume at the lower part of the stem. Conversely, *P. radiata* allocated relatively more volume than *E. grandis* at the upper 60% of the tree. Numerical integration of diameters in Figure 5.4 using the selected taper models yielded a total under-bark volume slightly larger for *E. grandis* (1.423 *vs* 1.383 m^3).

A summarised discussion and a listing or the key conclusions of all topics covered in Part I are presented in Chapters 13 and 14, respectively.

PART II

MODELLING STAND STRUCTURE AND DYNAMICS

CHAPTER 6

INTRODUCTION

6.1 BACKGROUND

Early growth models described forest development only by stand statistics such as mean top height, basal area or volume per unit area. This may be sufficient for some general management decisions, but it is widely acknowledged that modern forest planning requires more detailed information. This can be achieved either by (i) diameter distribution models; (ii) dis-aggregative models or (iii) individual-tree models. If these methods are to be included in an overall modelling system for a given population, the information at the different levels of resolution (i.e. stand, diameter class, and tree) should be compatible. That is, the sum of all tree basal areas in a plot should be equivalent to the plot estimate of basal area.

Most growth and yield modelling studies with compatible stand- and tree-level resolution have assumed that stand-level-based projections are more reliable than projections obtained by aggregation of predicted tree attributes. Therefore, tree-level-based estimations have been normally constrained so that they total the same as projected stand-level outputs (e.g. Woollons and Hayward 1985; Harrison and Daniels 1988; Candy 1997; Zhao 1999). One reason for that may be the ample body of knowledge on whole-stand dynamics. On the other hand, tree-level growth studies are scarcer and confined to the last few decades.

There are very few studies that have compared the overall performance of individual-tree and stand-level models (Knowe *et al.* 1997; Ritchie and Hann 1997a). Both studies concluded that stand-level projections obtained by aggregating projected individual-tree growth were as accurate or more so than stand-level projections themselves. Unfortunately, both studies have some limitations.

The dataset analysed by Knowe *et al.* (1997) consisted of only 10 plots (with 72 observations) from a spacing study that explored extreme densities (up to 13 889 trees/ha). The plots varied in size depending on the density, but most of them (seven out of ten) were extremely small (less than 0.02 ha). These limitations of the dataset restricted the reach of the conclusions severely.

The study by Ritchie and Hann (1997a) was conducted on *mostly* even aged Douglas-fir stands with minor components of other conifers. This stand structure may resemble to some extent even-aged single-species plantations, but it is not even-aged. The presence of other species and the existence of various tree ages probably made the structure and the dynamics of these stands more complex than they would be in plantations. It is usually acknowledged that individual tree models are best suited to represent complex stand structures (Knowe *et al.* 1997; Ritchie and Hann 1997a, 1997b). Additionally, the study by Ritchie and Hann (1997a) included only single 5-year-interval projections.

Another drawback of the two studies just mentioned is that they only considered diameter and basal area growth. The lack of height and, therefore, volume estimations make them somewhat incomplete.

During the course of the present study, another study comparing a tree-level and a stand-level growth models was published (McDill 2000). The models reported by McDill have a different scope to the ones developed in the present study. Firstly, McDill's models were developed for mixed natural forests. In addition, the stand level model consisted of a simple volume increment equation (static model as opposed to dynamic models, as explained in Section 7.3.1). Fitted statistics were very poor owing to data deficiencies, but the results indicated that the simpler stand-level model outperformed the more complex tree-level model (McDill 2000). One of McDill's concluding remarks was: *"To the best of my knowledge, this is the first time a tree-level growth model and a stand-*

level growth model have been developed and compared against each other using the same dataset". This remark, although not true (see Ritchie and Hann 1997a and Knowe *et al.* 1997), emphasises the scarcity of this type of study.

A thorough comparison of different approaches for developing growth and yield simulators with individual tree or diameter class resolution seems to be non-existent.

Results from the comparison of modelling approaches are likely to be data dependent. In order to reach more robust conclusions it would be necessary to use various datasets. Five or six datasets would be ideal, but considering the limited time available for this study, using three datasets was deemed reasonable. The species chosen, namely *Eucalyptus grandis*, *Pinus radiata* and *Pseudotsuga menziesii* (Douglas-fir), are all important plantation species being managed under rotations ranging from 10 (*E. grandis*) to 40 years or more (*P. menziesii*). Data for *P. radiata* and *P. menziesii* came from permanent sample plots (PSP) from several forests located in the Central North Island (CNI) of New Zealand, while *E. grandis* data come from PSPs in Uruguay.

For *P. radiata* growing in the CNI there are several stand level models, the more commonly used of them now being the PPM88 model (García 1988a). The output resolution of the PPM88 model can be enhanced by generating Weibull-based diameter distributions with procedures reported by Lawrence (1990). A nation-wide individual tree model for *P. radiata* has been completed during the course of the present study (Gordon and Shula 1999).

For *P. menziesii* growing in the CNI there is a whole-stand and a diameter distribution model developed by Xu (1990). A thorough evaluation of this model performed by M.O. Kimberly and R.L. Knowles (pers. comm.) in 1996 indicated that in a typical run from an age of about 30 years, the model over-estimated basal area increment by about 20%. At the time the present study was being undertaken, another stand level model, largely driven by live crown length per hectare, was also being developed (R.L. Knowles, pers. comm.). No individual tree models are publicly available for this species in New Zealand.

Stand-level models for both New Zealand grown species in the CNI were built more than 10 years ago. Therefore, the development of updated models incorporating new developments in growth and yield modelling, as well as enhanced datasets with a decade of new measurements was considered relevant.

E. grandis is one of the main plantation species in Uruguay. However, given the relatively short history of the forestry sector in Uruguay there are no growth and yield models available for this species or any other in the country.

6.2 STUDY OBJECTIVES

The objectives of this part of the study (Part II) can be categorised into methodological and practical. The former type of objective is concerned with the details of model construction and the exploration of best modelling techniques for each component. Practical objectives are related to the functional usefulness of the developed models.

6.2.1 Methodological objectives

The overall methodological objective of this study was to compare different modelling approaches with comparable output resolution for predicting stand structure and dynamics. These approaches included (i) a diameter distribution model; (ii) a relative-basal-area-based dis-aggregative approach and (iii) an individual tree model. In order for the three approaches to be comparable, tree-level outputs from the latter two approaches need to be grouped in diameter classes.

A second objective was to determine whether stand basal area and stocking can be estimated by stand-alone individual tree models as accurately as with stand-level models. The answer to this question can shed light on the dilemma of whether individual-tree models necessarily need to be adjusted by stand models or whether they can be built independently.

Other specific methodological objectives were:

- to compare various equation forms for modelling stand-level variables and identify those consistently superior for certain variables.

- to compare different approaches to diameter modelling (diameter-increment equations *versus* diameter difference equations) in terms of diameter estimation capabilities and usability for constructing individual-tree models.

- to compare different techniques for estimating individual-tree heights from diameter and mean top height measurements.

6.2.2 Practical objectives

From a practical standpoint the overall objective of this study was to develop growth and yield models providing increasing levels of output resolution for increasing level of detail in the input data. A minimum level of detail in the input data would include only stand-level statistics of basal area, stocking and mean top height. This information would be sufficient to perform stand-level projections. The next level of detail in the input data would include diameter statistics (i.e. maximum and variance or standard deviation) in addition to previous stand-level statistics. With this information it would be possible to perform projections with diameter-class output resolution. The most detailed level in the input data would be a tree list, which would allow projections to be performed with tree-level output resolution.

All these models should be compatible, meaning that the summation of basal areas from all trees should be the same as the stand-level estimate of basal area. The same rationale should apply to numbers of trees.

The developed models should be programmed in a user-friendly computer framework in order to allow users to take full advantage of the models' capabilities.

6.3 SCOPE OF THE STUDY

Models for the New Zealand grown species were developed with data from the Central North Island (CNI). As the models have not been evaluated for other regions, their use should be restricted to the CNI region.

The *E. grandis* model was built with data from Zones 7, 8 and 9 in Uruguay according to the CIDE (1967) soil classification. The database available for this species was very limited particularly for plots aged ten years or more, for plots with thinning treatments and for plots located in Zones 8 and 9. Given these data limitations, the model should be deemed provisional.

Preferably, input data for the models should be within the ranges in datasets used for model construction (see Table 8.2). Given the asymptotic and other logical features of the main components, limited extrapolations beyond these ranges can be safely performed.

6.4 NOTATION

Unless otherwise stated, the following symbols and definitions apply throughout this thesis:

$\alpha, \beta, \gamma, \delta, \theta$, a, b, c: parameters or regression coefficients in an equation

ALT: altitude above sea level (m)

SEE: standard error of parameter estimates

MSE: mean square error

t_1: age of tree or stand in years at the commencement of a projection interval

t_2: age of tree or stand in years at the end of a projection interval

tt: age of thinning in years

ha: hectare

G_i: stand basal area (m^2/ha) at stand age t_i

Gt: thinned basal area (m^2/ha)

Gb basal area before thinning (m^2/ha)

MTH_i: mean top height (m) at age t_i

h: tree height (m)

SI: site index, i.e. MTH at the reference age (10, 20 and 40 years for *E. grandis*, *P. radiata* and *P. menziesii*, respectively)

V_i: under bark stand volume (m^3/ha) at age t_i

N_i: number of stems per hectare (stocking) at age t_i

nn: starting stocking, i.e. number of stems per hectare at plot establishment or after thinning

n_i: number of trees in a diameter class or number of trees in a hectare represented by one tree (i.e. expansion factor)

n: number of observations in a dataset

BH: breast height (1.30 m for Uruguay and 1.40 m for New Zealand)

d or dbh or dbhob: tree diameter at BH outside bark (cm)

ΔD: annual increment in dbh (cm)

QMD: quadratic mean diameter (cm)

MD: arithmetic mean diameter (cm)

MTD: mean top diameter (cm)

D_{max}: maximum plot diameter (cm)

D_{std}: standard deviation of plot diameters (cm)

ln(x): natural (base e) logarithm of the variable x

$\log_{10}(x)$: base 10 logarithm of the variable x

Min (1%): first percentile of a variable's distribution

Max (99%): ninety-ninth percentile of a variable's distribution

X_i: i^{th} explanatory variable in a model

PSP: permanent sample plot

p.d.f.: probability density function (e.g. Weibull distribution)

DDM: diameter distribution model

ITM: unadjusted individual tree model

ITM_{adj}: adjusted individual tree model

RBA or R: relative basal area

BAL: basal area of all trees in a plot larger than or equal to the subject tree (m_2/ha)

GF: growth and form rating used to grade genetic improvement for *P. radiata*

CHAPTER 7

LITERATURE REVIEW

7.1 FOREST GROWTH AND YIELD MODELLING

A model is a simplification or an abstraction of a real system. In a forestry context, growth and yield models can be developed to represent or simulate the natural dynamics of trees, stands or forests. Given the long cycles involved in forestry, the availability of models for projecting the growth and yield of stands is essential, allowing us to make more sensible decisions than would otherwise be possible.

Given a set of stand or tree attributes (e.g. basal area, tree diameter, stocking and height) at a given age (t_1), growth and yield models can estimate the most likely values of these variables at a specified future age (t_2).

Applications of growth and yield models are numerous. For instance, given a collection of stands (forest) with various ages, heights, basal areas and stockings, it is possible to predict the future timber production of the different stands with a growth model. This allows rational management of the wood flows, through allowing sound decisions on rotation lengths to be made as well as the rate of new afforestation required to meet a certain target of annual timber supply.

Growth modelling and yield forecasting are indispensable tools for updating inventory information, management planning, evaluating silvicultural options and scheduling harvests. New applications of forest growth and yield models include carbon stock and biomass estimations for carbon credits trading and bio-energy projects.

The history of the development of growth modelling approaches and techniques from the early normal yield tables to the modern growth and yield models has been described elsewhere (e.g. Munro 1984; Xu 1990; Mason 1992; Temu 1992; Villanueva 1992; Goulding 1994). The following sections will focus on descriptions of more recent modelling approaches.

7.2 LEVEL OF RESOLUTION OF GROWTH AND YIELD MODELS

In the popular classification of forest growth models proposed by Munro (1974), the models are distinguished on the basis of two features: inter-tree dependency status and primary unit parameter requirements. Using these criteria, the following types of models have been identified:

- Distance-dependent tree-level models
- Distance-independent tree-level models
- Stand-level models

Distance-dependent tree-level models assume that the primary unit of stand modelling is the single tree and that inter-tree distance is a necessary parameter. Thus, each tree must be located in the model in terms of a spatial co-ordinate system (Munro 1984). Growth and mortality probabilities for each tree are expressed as functions of their dimensions and of the relative position and dimensions of their neighbours (García 1988).

In distance-independent tree-level models, the primary unit of stand modelling is also a single tree, but intertree distance is not a necessary parameter (Munro 1984). These models occupy an intermediate position between stand-level and tree-position models in

terms of state description detail and costs (García 1988b). Distance-independent tree-level models are discussed in Section 7.7.3.

Stand-level models describe the development of stand statistics such as basal area, mean top height and stems per hectare over time. In some situations, this type of model is likely to be appropriate for management planning of forest plantations (García 1988b). Depictions of size class distributions are likely to be necessary for management purposes. Size-class distributions affect the amounts of wood available for different end uses, harvesting costs per unit volume and the industrial conversion rate. Estimations of size distributions can be made from the projected values (output) of a stand-level growth model by a *diameter distribution model* (section 7.7.1). However, there can be advantages in retaining the individual-tree attributes measured at inventory and projecting them with individual tree models. As well as *dbh*, stem quality information such as pruned height, and defects such as fork height can be retained and thus give more accurate numerical descriptions and average dimensions of log products than those obtained from the diameter distribution approach (Candy 1997).

Vanclay (1994) identified an intermediate type of model, *size class models*, which is a compromise between whole stand models and single-tree models. Size class models provide information regarding the structure of the stand. Several techniques are available to model stand structure, but one of the most widely used is the stand table projection approach (e.g. Pienaar and Harrison 1988), which produces a histogram of stem diameters (Vanclay 1994).

Goelz (2000) challenged the traditional grouping of growth models into discrete classes. He stated that all growth and yield models could be viewed as diameter distribution models that merely differ in regard to which diameter distribution was employed and how the distribution was projected to future conditions. For instance, from a whole-stand model the implied diameter distribution would be a "spike of probability" at the average tree.

7.3 GROWTH MODELLING APPROACHES

Growth models and modelling approaches can be categorised in various ways according to different standpoints. These categorisations can be somewhat arbitrary and their purpose should be viewed just as a way of facilitating the description and understanding of the alternative approaches available. In reality, however, some models may not be easily labelled as belonging to one clearly bounded category.

7.3.1 Static *vs.* dynamic models

Depending on the method of predicting the quantities of interest (usually volume and/or mean diameter), growth models can be regarded as static or dynamic (García 1988b). A static model implies that the quantities of interest are directly predicted. This approach can give good results for unthinned stands, or for stands subject to a limited range of treatments, for which long-term experimental data are available (García 1988b). When the prediction of future volume yield at a chosen projection age is based on predicted values of stand statistics (or state variables) such as top height, number of surviving trees per unit area and basal area, the model is said to be dynamic (García 1988). This approach is the most commonly used in recent forest modelling studies.

7.3.2 Explicit *vs.* implicit (stand level) models

Clutter *et al.* (1983) distinguished two types of model according to the way they predict future yields. An *explicit* prediction of future yield implies direct calculation of stand volume by a suitable equation. This equation, in turn, includes the projected values of stand variables (e.g. mean top height and basal area) as in the dynamic models mentioned previously. A diameter distribution model, also known as an *implicit* prediction of future yield (Clutter *et al.* 1983), predicts the number of trees per unit area by diameter class using probability density functions. The mean height of each diameter class is estimated by regression on *dbh* and other stand variables. The volume of the average tree in each *dbh* class is then obtained with a suitable two dimensional tree volume equation using predicted *dbh* and tree height. By multiplying this volume by the number of trees per

hectare in each diameter class and adding up the total volumes of all diameter classes, an implicit estimation of the volume per hectare is obtained.

7.3.3 Process *vs.* empirical models

Another way to classify forest growth and yield models concerns the condition of the forest that is being modelled and the purpose of the model (Bruce and Wensel 1987). Clearly, there is a wider range of possibilities for characterising even-aged mono-specific stands. On the other hand, a mixed uneven-aged tropical forest requires more detailed data to be properly modelled, and the choice of modelling approaches is constrained. With respect to the purpose of the model, Bruce and Wensel (1987) made a distinction between *process* and *empirical* models. Some process models, for instance, attempt to model the biological processes involved in the transformation of CO_2, nutrients and water into plant biomass through photosynthesis. They may explicitly consider precipitation, hours of sunlight and other environmental factors. Although these models help to provide a better understanding of growth and stand dynamics, few have become incorporated into forest management systems (Battaglia and Sands 1998). A good example of a simple process-based model of forest productivity is the one reported by Landsberg and Waring (1997). This model showed excellent correspondence between stand growth measurements and simulated stem growth over 30 years.

Empirical models are constructed using periodic tree measurement data, and as such, do not consider basic processes affecting tree or stand growth. This does not mean that empirical models do not provide biologically realistic predictions, nor does it mean that they are inferior to supposedly biologically-based models (West 1997). Empirical models may sacrifice specific details of growth processes in order to achieve greater efficiency and accuracy in providing information for forest management. Process models have also been termed as *models for understanding*, whereas empirical models are regarded as *models for prediction* (Bunnell 1989). The present study is solely concerned with empirical models, and is aimed at developing models suited to forest management and planning.

An intermediate approach between empirical and process-based models, the so-called hybrid models has been proposed to fine-tune traditional empirical models through *growth indices* reflecting changing conditions for growth (Snowdon *et al.* 1998; Snowdon *et al.* 1999; Snowdon 2000). This newer intermediate approach seems to combine the best features of empirical and process-based models. A good example of the utility of this approach is for annual inventory updating; the inclusion of actual climatic conditions can significantly improve the accuracy of predictions (Snowdon 2000).

7.3.4 Deterministic *vs.* stochastic models

Irrespective of its level of detail or the way it was constructed, a model can be deterministic or stochastic (Vanclay 1994). Given the same initial conditions, a *deterministic* model will always predict the same result. A *stochastic* model attempts to illustrate the natural randomness of prevailing environmental conditions by providing different growth predictions, each with a specific probability of occurrence. Deterministic and stochastic models have different purposes. The former type is effective for determining the expected yield and may be used to indicate the full potential of a stand, whereas the latter type may indicate the reliability of these predictions (Vanclay 1994). In practice, most of the information needs for forest planners and managers can be provided more efficiently with deterministic models (Vanclay 1994). Therefore, this is the approach to be adopted in the present study.

7.3.5 Clutter's approach *vs.* García's approach

Regarding the procedures used in the process of model development, a discrimination can be made between models in which the main state variables are simultaneously fitted (García 1984, 1988b; Law 1990; García 1994) and models in which each variable is independently fitted (Clutter *et al.* 1983). In García's approach, all state variables must be modelled with the same functional form of a particular model. On the other hand, different models or variants of the same model can be adjusted for each variable in Clutter's approach. This gives flexibility to choose the functional form that best fits each variable. Stand development is modelled over time in Clutter's approach, whereas stand top height is used in García's approach as a surrogate of time.

García's approach has been referred to as the *state-space approach* (García 1994). However, the key concepts attributed to the aforementioned approach, namely that the system is described by a number of state variables the rate of change of which is then modelled, also apply to the conventional modelling approach.

Clutter's approach has been the most widely used by forest modellers throughout the world. This approach is more flexible, parsimonious and less demanding in terms of data quality. Therefore it will be used in the present study.

7.4 SITE QUALITY AND SITE INDEX

Clutter *et al.* (1983) defined site quality as "the timber production potential of a site for a particular species or forest type". Site quality can be determined by a number of soil (fertility, depth, texture), climatic (rainfall, temperatures) and biotic (e.g. diseases, understorey vegetation) factors that influence the growth of a given tree species. In developing growth and yield models it is desirable and generally necessary to include a measure of site quality. Indeed, for a given species and region, the shape of the growth trajectories, the maximum attainable yield (asymptotic yield), or both, are likely to be affected by the site quality.

Methods for classifying site quality can be direct or indirect (Clutter *et al.* 1983) as briefly described below.

7.4.1 Indirect methods for evaluating site quality

When the species (or forest type) of interest is not present on the land area under evaluation, it is possible to use indicator species to ascertain the potential of the site. Other forest species are preferred over herbaceous or shrub species as the latter reflects only the fertility of the topmost soil horizons and not the deeper soil layers that are more relevant to tree species. Observed height growth of one tree species has been used to estimate the site quality for another tree species by Coile (1948) and Olson and Della-

Bianca (1959). Ure (1950) related forest site productivity in Kaingaroa Forest to the presence or absence of certain understorey plant species.

Using environmental site attributes is another possible way of predicting site quality for a given forest species. For instance, site index (see Section 7.4.2 below) of radiata pine in the North Island of New Zealand has been related to mean annual rainfall, temperature, soil nutrients, topsoil depth, soil penetrability and soil pH (Hunter and Gibson 1984). For a particular forest of that region (Kaingaroa), a simpler yet robust linear relationship between site index and altitude was proposed (Mountfort 1979).

Many studies have been implemented in order to find soil variables that better explain (or predict) the growth of *E. grandis* (De Moraes Goncalves *et al.* 1990; Noble *et al.* 1991; Louw 1997). Within physiographic areas suitable for the species, and excluding sites with extremely high deficiencies of the main nutrients, most of these studies have concluded that one or two variables influencing root development and water storage (typically soil depth), are the most influential on growth of the species.

For *E. maidenii* in Rwanda, 14 meteorological and A-horizon soil variables were analysed and the available moisture retention was the one which most affected the site potential for the species, explaining 33.6% of the site index variation (Gasana and Loewenstein 1984).

7.4.2 Direct methods for evaluating site quality

The direct assessment of the growth of the species of interest in a given site is usually a more accurate way of determining site quality potential[1]. Clutter *et al.* (1983) classified direct methods for evaluating site quality according to the kind of data used as follows: (i) from historical yield records, (ii) from stand volume data, and (iii) from stand height data. The first category has the limitation that for many areas there are no previous records from previous crops. Moreover, when these records do exist it is likely that the silviculture used in the past was different to that used for the current stand. The second

[1] Site index is generally associated with *site* quality, although it should be *crop* quality as different planting stocks may result in different site index estimates for a given site (Dr. A.G.D. Whyte, pers. comm.).

category has similar limitations as the first, as volume yield is strongly affected by stand density and other silvicultural operations.

The use of stand height data for representing site quality is based on the concept of *site index*. Site index is the average height of the dominant and co-dominant trees at a specified age called index age or base age. The index age is commonly selected to lie close to the average rotation age (Clutter *et al.* 1983). For *E. grandis* in Uruguay the base age is 10 (Sorrentino 1991, 1992, 1993), while for New Zealand grown *P. radiata* and Douglas-fir the base ages are 20 and 40 respectively (Burkhart and Tennent 1977a, 1977b).

As the definition of dominant and co-dominant trees can be subjective, it is now common to standardise it by using the 100 trees of largest *dbh* per hectare. In theory, dominant stand height is slightly affected by varying density levels, relatively stable under varying thinning intensities (unless thinning from above), and strongly correlated with stand volume. The mentioned invariance of dominant stand height with stand density (i.e. stocking) generally holds for plantation species under the most common range of stockings. However, it has been found that very low stockings can reduce dominant height growth of *P. radiata* in New Zealand (Maclaren *et al.* 1995). Mason (1992) pointed out other possible limitations of using site index as a measure of site quality, namely:

(i) environmental factors influencing height growth are not specifically included in dominant height models, leading to limited understanding and inaccuracy if these factors change (e.g. appearance or disappearance of diseases, increased atmospheric carbon levels);

(ii) height is usually measured with less precision than other stand variables as basal area; and

(iii) variations in height growth may not always be directly related to variations in volume growth if basal area growth is differently influenced by factors affecting growth.

Some growth modellers accounted for the third limitation by allowing the user to make basal area growth adjustments within a given site index (West *et al.* 1987).

Despite the mentioned shortcomings, site index is by far the most common tool used to represent site quality. This is because site index is relatively easy to derive and because of the lack of other feasible and sound methods. Moreover, it provides a simple numerical value that can be easily measured and understood by forest managers.

7.4.3 Development of site index curves

A stand-dominant-height *versus* age curve (site index curve) is required to compute site indices by reading the dominant height value at a specified base age. Site index curves can be classified according to the nature of the height/age curve families they generate as being (i) anamorphic, (ii) polymorphic-disjoint, or (iii) polymorphic non-disjoint (Clutter *et al.* 1983; Borders *et al.* 1984).

In an anamorphic curve family, the value of the dependent variable (y) of one curve at any age is a constant proportion of the y value of the other at the same age. In a polymorphic curve family, this proportionality relationship does not hold (Clutter *et al.* 1983) and the shape of the curves is different. Whilst anamorphic curves have always different asymptotes, polymorphic ones can have a common asymptote or varying asymptotes (Cieszewski and Bailey 2000). For polymorphic curves, the distinction between disjoint and non-disjoint is concerned with whether the curves cross each other (non-disjoint) or they do not (disjoint) within the range of interest.

Site index curves can be obtained by different methods, depending on the kind of data available. The method of construction, in turn, may restrict the type of site index curves that can be developed. The three methods identified by Clutter *et al.* (1983) are briefly described below.

7.4.3.1 Guide curve method

This method can be used to generate anamorphic site index equations from yield-type data (as opposed to re-measurement data) of stand height-age. Height-age data obtained from temporary plots covering a broad range of ages and sites need to be collected. Then, a number of yield equations of the form H=f(t, α, β, γ) can be fitted to the data (where H is stand dominant height, t is stand age and α, β and γ are parameters to be obtained by least-squares methods). The equation exhibiting the best fit within those equations conforming with expected behaviour with regard to upper asymptotes or inflection points should be selected.

The equation so fitted describes the *guide curve*, which represents the *average* height-age trajectory. Parallel curves can be drawn by proportionally varying the height value at each age, so that the whole range of data can be contained within the anamorphic curves.

For *Eucalyptus grandis* in Uruguay this was the method used to develop the existing site index curves (Sorrentino 1992). For the same species in the neighbouring province of Entre Ríos (Argentina), Glade (1999) also developed site index curves using the guide curve method.

7.4.3.2 Difference equation method

Data from permanent sample plots (PSP) or stem analysis are required for this method. The procedure is flexible and can be applied to any height-age model to produce anamorphic or polymorphic curve families (Clutter *et al.* 1983). By using the algebraic difference approach (Bailey and Clutter 1974) a yield model can be transformed in a difference equation. For instance, let us consider the following Chapman-Richards model:

$$H = A\left(1 - e^{-kt}\right)^{\frac{1}{1-m}} \qquad (7.1)$$

where A is the *asymptote*, k is a *growth rate* related parameter, and m is a *shape* parameter. At time t_1 and t_2 mean top heights are, respectively:

$$H_1 = A\left(1-e^{-kt_1}\right)^{\frac{1}{1-m}} \quad \text{and} \quad H_2 = A\left(1-e^{-kt_2}\right)^{\frac{1}{1-m}} \tag{7.2}$$

By making A the subject in the previous equations, equating the resulting equations and re-arranging, we get the following anamorphic equation:

$$H_2 = H_1\left(\frac{1-e^{-kt_2}}{1-e^{-kt_1}}\right)^{\frac{1}{1-m}} \tag{7.3}$$

This anamorphic equation was the selected site index curve for the current Douglas-fir growth model for the Central North Island of New Zealand (Xu 1990).

Analogously, by making k the subject we get the following polymorphic equation, which is the most common version of Chapman-Richard's model found in the literature (sometimes 1-m is denoted as c). The PPM88 model for *P. radiata* in the Central North Island of New Zealand contains the equation below as its stand height model.

$$H_2 = A\left\{1-\left[1-\left(\frac{H_1}{A}\right)^{1-m}\right]^{\frac{t_2}{t_1}}\right\}^{\frac{1}{1-m}} \tag{7.4}$$

Another polymorphic equation which can be derived by letting m be the free parameter is:

$$H_2 = A\left(\frac{H_1}{A}\right)^{\frac{\ln(1-\exp(-kt_2))}{\ln(1-\exp(-kt_1))}} \tag{7.5}$$

The latter form was found best for *Eucalyptus* plantations in Portugal (Amaro *et al.* 1998). It accommodates the possibility of having no lower inflection point, which is

likely to occur when the initial growth is rapid and the instantaneous growth rate is monotonically decreasing from time zero onward.

When t_2 is set to the base age in the previous three equations then the resulting H_2 will be an explicit estimate of site index.

7.4.3.3 Parameter prediction method

As the previous method, this procedure requires remeasurement data from permanent sample plots or stem analysis. The basic procedure involves the following three steps (Clutter *et al.* 1983):

(i) fitting a linear or non-linear height-age model to the data on a plot-by-plot (PSP) or tree-by-tree (stem analysis) basis;

(ii) assigning a site index value to each tree/plot using each fitter curve;

(iii) regressing the parameters of the fitted curves on site index.

This procedure is less common than the difference equation approach and it suffers from certain limitations that do not apply to the latter (see Clutter *et al.* 1983, pages 55-56).

7.5 DATA COLLECTION AND PREPARATION FOR GROWTH MODELLING

7.5.1 Introduction

Forest inventory can serve many purposes, but different procedures are required to satisfy various needs of different data users in an efficient way. In this respect, Vanclay (1994) identified the following information requirements and corresponding sample plot procedures:

(i) Resource inventory. Many sampling units are required, preferably covering a broad range of environmental and silvicultural gradients. Temporary inventory sampling units are usually the most cost-effective approach for resource inventory.

(ii) Continuous forestry inventory (CFI). Forest growth and harvesting can be monitored by remeasuring permanent plots either all of them or by sampling with partial replacement (SPR). Plots should be located in representative areas of the various forest types and stand conditions in proportion to their area.

(iii) Growth modelling. The required quality of data for growth modelling demands more rigorous procedures for establishing and measuring permanent plots, which are not necessary for CFI systems. Specifically, individual trees should be unambiguously identified, plots should be homogeneous and extremes of site and stand conditions adequately represented. In contrast to CFI plots, it is not necessary for the growth modelling plots to be numerically proportional to forest areas.

(iv) Long term monitoring of environmental change. Permanent plots aimed at providing information for this purpose would need to be measured with more level of detail than permanent sample plots for growth modelling. Cost considerations usually invalidate the generalised use of this plot type.

Only the third category is within the scope of this study.

7.5.2 Stem analysis

Although permanent sample plots are the most common data source for growth modelling studies, the use of stem analysis can be another valid alternative in some circumstances (Jerram 1939; Assmann 1970). For this method to be applicable, tree species need to exhibit clear annual ring patterns. In this sense, softwoods under colder climates are much more suited than hardwoods under warmer climates. For radiata pine, Wilcox (1987) proposed a retrospective sampling technique based on stem analysis

where individual stems can provide a complete series of height and diameter data. The method of tree ring analysis is more precise in older stands and for short-interval projections (García 1992).

For the construction of complex and accurate growth and yield models, stem analysis tools may not be adequate and permanent sample plots are generally preferred.

7.5.3 Permanent sample plots (PSP)

In setting up and maintaining a permanent plot system for growth modelling, the following factors should be considered (Vanclay 1994):

(i) Temporal distribution. To avoid year to year effects on growth, growth studies should be based on data collected over a long period. Additionally, it is desirable to conduct all measurements at about the same time of the year to minimise the introduction of bias from seasonal effects.

(ii) Spatial distribution. Under the premise that is better to interpolate than to extrapolate, plots should sample a broad range of locations (i.e. latitude, altitude, aspect, *et cetera*). A systematic grid sample may provide a good coverage of various conditions, although some sort of stratification may also be convenient.

(iii) Site factors. Plots should be sited throughout the whole range of site qualities for forest growth.

(iv) Stand conditions. As growth is strongly affected by stand density, it is recommended that the permanent plot system include a variety of stand conditions. Experiments or clusters of plots including extreme stockings are desirable.

To define approximately the number of plots and the interval between successive measurements, the previous factors along with the level of sensitivity and precision expected from the models and the availability of resources have to be jointly considered.

Minimum standards for establishing and measuring permanent sample plots for plantations in New Zealand have been thoroughly discussed by Ellis and Dunlop (1994) and Ellis and Hayes (1997). Salient points from these standards which should be mentioned are: (i) circular plots are preferred, (ii) the plot size should be defined to ensure that at least 20 live trees can be measured at the final density, (iii) plot area should be corrected when slope is higher than 10%, (iv) plot boundaries and individual trees should be clearly and unambiguously marked. Circular plots may be difficult to locate because they are less clearly delineated that square or rectangular plots.

7.5.4 Growth intervals

Data used for growth modelling can be used in two basic forms, namely yield form and interval form. Yield-form data are useful for developing initial growth models capable of simulating stand conditions prior to pruning and thinning, according to varying site and silvicultural situations during the establishment phase (e.g. Mason 1992; Mason and Whyte 1997; Zhao 1999). In addition, yield-type data allow a better detection of site and genetic effects on tree growth than interval-type data do.

Growth and yield models for older ages are generally obtained from real growth series (i.e. interval-type data) derived from PSP measurements, although yield data may be appropriate for developing simpler models for populations with homogeneous site and stand conditions. Model coefficients can be easily obtained by fitting difference equations typically obtained from an algebraic difference approach (see Section 7.4.3.2). The generic form of difference equations is $Y_2=f(Y_1, t_1, t_2)$, which requires PSP data to be arranged in interval form.

When a PSP has been measured on two occasions, only one interval (t_2-t_1) can be formed. If a third measurement is recorded (i.e. t_3), then not only the interval t_3-t_2 can be added, but also the interval t_3-t_1. The inclusion of the latter interval is not common and represents a relatively recent development in growth modelling (Borders *et al.* 1987). From n PSP measurements a number of $\binom{2}{n}$ growth intervals can be constructed, as

opposed to n-1 in the conventional non-overlapping data structure. This massive multiplication of the potential growth intervals is *a priori* very appealing from the modellers' perspective.

The use of an all-possible-interval data structure allows the use of long projection intervals, which could potentially improve the model's capability over long projections. This advantage would more likely accrue provided that the developed models are path invariant (see Section 7.6.2). Implications of the path-invariance property on the way that intervals should be constructed have received little attention in the literature.

Borders *et al.* (1987) investigated the model performance of three methods of data preparation, namely non-overlapping growth intervals, all possible intervals, and longest interval only. They concluded that the relative merits of each data structure depended on the model form. It might be possible that the implications of using different data structures can be data-dependent, as well. Lee (1998) and Zhao (1999) compared different data structures for modelling stand development of Douglas-fir and radiata pine from large datasets and both concluded that the use of mixed intervals (all possible intervals) led to the best results. Amaro *et al.* (1998) compared a non-overlapping data structure against an all-possible-interval data structure for modelling mean top height and concluded that the model constructed with the latter data structure performed slightly better.

The use of all possible intervals has the disadvantage that the amount of autocorrelation in the data is increased (Borders *et al.* 1987). In order to minimise this shortcoming yet retain the positive aspects of the method, a proper interval sub-sampling could be a sensible solution.

7.6 METHODOLOGICAL CONSIDERATIONS ON THE CONSTRUCTION OF GROWTH MODELS

7.6.1 Model components

For stand-level growth and yield models it is common and generally sufficient to describe the current state of the stand by three variables: dominant height, basal area and stocking. From these three variables, the future state of the stand can be predicted with reasonable accuracy (García 1988). More sensitivity and accuracy can be achieved if the description of the initial state is augmented with environmental variables such as altitude (e.g. Mason 1992; Zhao 1999) or geographic region (e.g. Temu 1992; Lee 1998). The description of the initial state or its changes can also accommodate the effect of silvicultural operations. For example, West *et al.* (1982; 1987) developed a growth model (EARLY) that is sensitive to changes in canopy structure derived from pruning and thinning operations. Although more refinements are possible with greater data quality and quantity, there is a potential risk of losing robustness and consistency if the description of the initial state becomes too cumbersome. Moreover, when the model is implemented, an excessive level of required inputs might prevent the intended user running the model at all if some required information is missing.

Stand-level models may provide volume estimations as well. These are usually obtained from basal area and height estimates (e.g. Woollons and Hayward 1985) and, sometimes, from predicted stocking (García 1984) and age (Amateis *et al.* 1986).

When stand level estimates are dis-aggregated into *dbh* class information (diameter distribution models) at least two other components are required. One of these extra components is for estimating the variance of the diameters while the other is for predicting the extreme diameter (maximum or minimum). With these extra components it is possible to solve the parameters of a probability density function (e.g. Weibull).

In comparison to stand models, individual-tree models or stand table projection approaches require a more detailed description of the initial state. A tree list or a stand table (number of trees by diameter class) is required for this type of model. At least three

components are required for a complete individual tree model, namely a diameter increment model, a model for estimating the probability of tree survival and an individual-tree-height model. More detailed individual tree models also require extra environmental information such as rainfall, altitude and fertility indices (Gordon and Lawrence 1994; Shula 1997b).

7.6.2 Functions used

Numerous sigmoid models have been proposed for modelling plant growth. These include the log-reciprocal (Schumacher 1939; Clutter 1963), the so-called Chapman-Richards model (Von Bertalanffy 1957; Richards 1959; Pienaar and Turnbull 1973), Weibull model (Yang *et al.* 1978), Hossfeld model (Woollons *et al.* 1990) and others listed in Appendix 1.

From the yield form of all these models it is possible to obtain difference equations or projection equations, through the algebraic difference approach (Bailey and Clutter 1974). Clutter *et al.* (1983) noted four desirable properties of equations used for growth and yield modelling, namely:

(i) *Compatibility* of growth and yield (i.e. integral-derivative relationship of growth and yield).

(ii) *Consistency* in the projections. No change for zero elapsed time; as t_2 approaches t_1, Y_2 should approach Y_1.

(iii) *Path-invariance*. The estimate at a given final age should be the same irrespective of the number of projections used to get it; that is predicting Y_3 from Y_1 in a single step should yield the same value as predicting Y_2 from Y_1 and then Y_3 from Y_2.

(iv) *Asymptote*. As t_2 approaches ∞ Y_2 should approach an upper asymptote.

7.6.3 Modelling stand-level mortality

A number of projection equations have been used by growth modellers to model the development of the number of live stems per unit area over time (see for example Clutter *et al.* 1983). These equations are usually inverse-sigmoid-shaped and always predict decreasing stocking over time. In reality, though, some plots show no mortality in certain intervals. If these intervals were not used for model development, mortality would be overestimated. On the other hand, including such intervals would lead to convergence problems when fitting non-linear regressions and ill-fitting residual plots, which may complicate the comparison and evaluation of candidate models (Woollons 1998). A theoretical solution to this problem, consisting on a two-step regression process, was proposed by Woollons (1998). The first step is a logistic regression that predicts the probability of stand mortality occurring. In the second step, a stocking projection equation is fitted only to data exhibiting mortality. The final stocking estimate is obtained by combining the previous two components as follows:

$$N_{adj2} = N_1 - p\left(N_1 - \hat{N}_2\right) \tag{7.6}$$

where N_{adj2}= adjusted live stems/ha at time t_2; N_1 = live stems/ha at the commencement of the period; $\hat{N}_2$ = live stems/ha at time t_2, estimated by the stocking projection equation fitted to plots exhibiting mortality; p = probability of stand mortality occurring (from the logistic regression).

For modelling the probability of stand mortality occurring, Woollons recommended using data with a constant period between measurements to avoid bias in predictions. This would not be suited to data structures containing varying intervals (especially when longer intervals are included in addition to non-overlapping intervals). To overcome this problem, the period (or interval length) may be included in the logistic model, as it is clear that the probability of stand mortality occurring will be higher over longer intervals.

Apart from its theoretical merits, the practical implications of Woollons' two-step approach for modelling stand mortality should be considered. Comparisons of stocking

estimates obtained with the two approaches (i.e. conventional approach and Woollons' approach) against actual values would clarify this aspect.

7.6.4 Parameter estimation

The most commonly used statistical technique to estimate the parameters of non-linear (sigmoid) functions used in forest modelling is non-linear regression. This is an iterative procedure, which estimates the set of parameters that minimises the sum of the squared residuals. A set of starting values for the parameters should be specified by the modeller.

Ordinary least-squares (OLS) procedures are based on a number of theoretical assumptions. Let f represent any linear or non-linear function, $\mathbf{X}$ the vector of explanatory variables (e.g. Y_1, T_1 and T_2), and ϕ the vector of parameters (α, β, γ, etc.):

$$Y_2 = f(\mathbf{X}, \phi) + \varepsilon$$

Statistical procedures to estimate ϕ assume that the error term, ε, is normally distributed with zero mean, constant variance, and independence between cases.

Repeated measurements from PSP data are correlated, as the value of a variable at the time of a measurement subsumes the values at previous measurements. When the data used in the regressions are increments (rather than actual data of the variables), the level of correlation between measurements is lower (García 1984).

The assumption of independence between explanatory variables is also usually violated. Dependent variables such as height, basal area and stocking are often correlated (García 1984). This statistical violation of the least-squares assumptions is called *multicollinearity*.

Violation of least-squares assumptions results in underestimation in the standard error of parameters, which in turn produce unreliable significance tests. To avoid these problems, García (1984) proposed an alternative modelling approach (see Section 7.3.5). It consists

of simultaneously estimating the parameters for all state variables of the modelling system by means of stochastic maximum likelihood.

The practical importance of the statistical problems associated with the use of least-squares methods with PSP data has been often challenged. It is generally accepted that, despite underestimation of error variance, parameter estimation is unbiased (e.g. West 1995). Sullivan and Clutter (1972) compared least-squares estimation with maximum likelihood for growth modelling and found no differences of practical importance in the values of estimated parameters. Mixed effects analysis techniques are available for estimating parameters from autocorrelated data (see Section 1.3).

If the practical and biological reasons for building the models prevail over the strict statistical considerations, the criteria of minimising the residual sum of squares of the non-linear regression can be initially applied for parameter estimation. Then, hypothesis testing of parameters' significance can be performed with a sub-sample of the dataset containing autocorrelation-free data. There are many ways to check whether the developed model is unbiased and whether its errors are normally distributed. These include graphical and numerical methods as explained in Section 7.6.8.

7.6.5 Localising growth models

There has been some contention over the geographical extent to which growth models can be applied. On the one hand, it can be proposed that that the use of specific models for specific localities is required, to account for the particularities of growth patterns for each locality. On the other hand, it can be argued that such stratification may lead to the development of an excessive amount of models, each of them fitted to less complete databases than it would be possible by their grouping. Whyte *et al.* (1992) addressed this topic and suggested that excessive specificity in model coverage is not usually warranted, statistically or biologically. Moreover, they argued that such over-stratification would not be advisable from an operational standpoint.

If the development and use of more general growth models are to be advocated, one could raise the question of whether this would reduce the sensitivity, accuracy and

precision of the models. Many studies have shown that regional sensitivity can be granted by tailoring the projections of an overall model to local variations. This can be achieved by various approaches.

One way to localise predictions from a regional growth model to local conditions is the simple means ratio. It consists of multiplying the overall estimate by the ratio of the average locality prediction to the average overall prediction. Smith (1983) used this approach to adjust annual diameter growth estimates from the STEMS regional model (Shifley and Fairweather 1983).

A more sound and popular approach is the use of binary indicator variables (dummy variables). In statistical terms, the use of dummy variables can be viewed as formulating an analysis of covariance (ANCOVA), where the dummy variable is a covariate incorporated into the original regression. Dummy variables can adopt only the values of 0 or 1 and therefore they activate or deactivate their associated parameters. They were utilised to model binary effects such as weed control/no weed control (Mason 1992), or whether the simulation applies for a particular locality or not (e.g. Temu 1992, Lee 1998). The general approach has been demonstrated elsewhere (Gujarat 1970) but can be briefly outlined as follows. Let α and β be the fitted parameters of a linear or non-linear regression. Suppose the regression was fitted to data from two strata (e.g. two locations or two treatments). To test the null hypothesis that an overall model is sufficient (the alternative hypothesis would be that two models are required) the parameters α and β can be substituted by the expressions $(\alpha_0+\alpha_1 X)$ and $(\beta_0+\beta_1 X)$ respectively. The dummy variable X is set as 1 if one of the treatments applies and 0 if it does not. Note that for n strata only n-1 dummy variables are required as one stratum can be set as the default. If the parameters α_1 or β_1 are significantly different to zero (either one or both) it means that the adjustment provided by the parameters associated with the dummy variables are justified.

A much more complex approach for localising growth models, based on Bayesian methods, has also been proposed (Berkey 1982). This method requires high statistical knowledge, special algorithms for solving the parameters, and can only be applied when the assumptions of Bayes' theory are met. The increased complexity of this method may

not be justified in some cases, especially when the quality of the information is not ideal (Temu 1992).

7.6.6 Incorporating silvicultural treatments

In the previous section, the possibility of localising growth and yield models according to dichotomous-type treatments was mentioned. In reality, though, silvicultural treatments are not always dichotomous but they can consist of a continuum of doses (e.g. fertiliser) or intensities (e.g. thinning). Therefore, the level or intensity of the treatment should be specified in some cases.

The effect of a given treatment is likely to vary with the age at which the treatment was applied (t_t) and the elapsed time between the treatment and the projection age (t_2-t_t). These two factors can be introduced as linear modifiers of model parameters. However, the inclusion of t_2 usually implies that the path invariance of the model is lost. More complex modifications are therefore required.

Previous studies incorporated thinning terms into basal area projection equations (Bailey and Ware 1983; Murphy and Farrar 1988). These terms included a numerical variable describing the type of thinning along with the time of thinning (t_t) and the projection age (t_2). In both cases the thinning effect on basal area growth was lower for delayed thinnings and diminished as the elapsed time after thinning increased. However, it was unclear from the previous studies how repeated thinning would be accommodated in the implementation of the model. This raises the question of whether or not the effect of successive thinnings is cumulative or whether or not only the effect of the last thinning should be considered.

Fertilising effects were also included into basal area and stand height models by Martin *et al.* (1999). In the mentioned study, a variation of the Pienaar and Rheney (1995) approach to incorporating responses to silvicultural practices was used. This adjusted difference-form model (the adjustment proposed by Pienaar and Rheney was in a yield form) allowed the fertilising effect to be sensitive to the age of treatment and the elapsed time after treatment, while maintaining the property of path invariance. Moreover, it

helped alleviate any overprediction due to double accounting of the fertilisation response (Martin *et al.* 1999). This was accomplished by first subtracting the effect of the treatment on the initial basal area (i.e. adjusting back to the unfertilised baseline), and then adding the fertilising response at projection age to the predicted unfertilised figure.

7.6.7 Enhancing model projection ability

In this section, a few general principles that should be applied when developing growth and yield models are outlined. Some of them were discussed in previous sections and are simply summarised or emphasised here.

A key aspect of the modelling process is undoubtedly the quality and quantity of data available. Ideally, the dataset should have a wide coverage of the target population, spatially and temporally (Vanclay 1994; Vanclay 1995). Sampling errors from inappropriate or very limited databases may cause serious estimation bias, increasing the risk of having confounded factors. It is also desirable that the database contains not only passive monitoring plots (typical stands), but also experimental plots subject to more extreme silvicultural treatments, such as very severe thinnings (Burkhart 1987; Vanclay *et al.* 1995). The latter type of plots will confer the model more sensitivity for comparing a variety of tending regimes.

Once a given dataset is available, the next step is the construction of projection intervals. As discussed in Section 7.5.4, projection intervals may include non-overlapping intervals only, or a mixture of them with longer intervals. Some studies indicated that the inclusion of longer intervals could enhance model projection ability, particularly for long projections (Lee 1998; Zhao 1999). Other studies indicated that the results are likely to depend on the functional form and data source. However, it has never been proved that using only non-overlapping intervals can lead to better model capabilities.

If the data are severely biased towards a particular type of site quality, tending regime, age or period between successive measurements, it may be appropriate to sub-sample the data to balance its coverage. If this were done, some data would be left out of the final fitting dataset.

Many equation forms and regression procedures are available. Some equations tend to fit consistently better for certain variables (e.g. Schumacher equation for basal area per unit area). However, the relative goodness-of-fit of several equations is usually data-dependent (see for example Woollons and Wood 1992) and, thus, it is always convenient to try as many equations as possible for each variable, and select the one exhibiting the best fit.

Parameter estimates in non-linear regressions are obtained by an iterative procedure (Section 7.6.4). Care should be taken that the procedure has not converged to a local minimum as opposed to the global minimum of the sum of squared errors. A simple way to check for a global minimum solution is to (i) re-run PROC NLIN with a different method (Gauss-Newton, Marquardt, DUD) and (ii) re-run PROC NLIN with different starting values (Staudhammer and LeMay 2000).

Once the best equation for a given variable has been selected, the model can be fine-tuned by adding additional explanatory variables. The selection of candidate variables should be based on established silvicultural principles or in the inspection of plots of residuals against independent variables. Dummy variables for localities or binary-type treatments can also be considered as potential adjusters. Additional variables can enter the model modifying the original parameters in a linear or non-linear fashion. In order to maintain the path-invariance property of the models, any extra variable introduced into the model should be a constant (e.g. site index, altitude). More complex modifications of a model form which keep the path-invariance property have also been proposed (Bailey and Ware 1983).

When adding explanatory variables, care should be exercised to avoid creating unrealistic effects when the model is applied to independent data. This problem may occur if the added variable is confounded with other effects (e.g. site, age) in the database used for model fitting.

7.6.8 Model evaluation and validation

Although the terms model *evaluation* and model *validation* sometimes are used interchangeably, there is a tendency to keep the latter for situations where the model is tested against data that were not used for model development (i.e. independent data). Applying the model to an independent dataset is the safest way to ascertain its projection capabilities. However, this is not always possible as setting aside validation data reduces the amount of available data for model construction. The criteria to define independence of data are not always clear. At one extreme data coming from only a different source or region can be considered independent. Other possibilities include the use of plots sub-sampled from the original dataset or even sub-sampling individual measurements within plots. Irrespective of the independence level of data, established procedures for evaluating forest growth and yield models are available.

A clear distinction can be made between *qualitative* and *quantitative* procedures for evaluating model performance (Vanclay 1994; Soares *et al.* 1995; Gadow and Hui 1999). Qualitative evaluation or *model criticism* (Vanclay 1994) should comprise a critical appraisal of theoretical and biological aspects of the model. It should establish whether the whole system is properly integrated, and whether the outputs from a variety of initial conditions agree with sound biological and silvicultural principles. For example, the model should forecast lower total volume yields and greater average tree diameters for intensively thinned stands than for unthinned stands.

Quantitative procedures include the analysis of residuals (i.e. observed minus predicted values) through graphical and statistical tests. There are many aspects of the residuals that serve as model qualifiers. The Mean Residual (MR) is an important measure to be considered. Obviously, the closer MR is to zero, the better. The sign and magnitude of MR could be an initial indicator of the overall level of over/underestimation of the model (*accuracy*). However, a low mean residual does not ensure that the model is *precise,* or that it will provide unbiased projections over a range of values of the independent variables. Even an extremely flawed model can have a zero MR, as long as positive deviations cancel out negative deviations.

Theoretically, residuals should be normally distributed (around a zero mean). There are many ways, either graphical or through statistical tests, to evaluate the distribution of the residuals. Graphical methods include visual inspection of residual frequency histograms, the normal probability plot (e.g. Woollons 1998), and the cumulative percentage of the residuals compared to that in a perfect normal distribution. Statistical tests to check for normality include the Shapiro-Wilk and the Kolmogorov tests (SAS Institute Inc. 1990).

Skewness and kurtosis are numerical indicators commonly used to evaluate distributions of residuals. Skewness measures the tendency of the residual deviations from their mean to be larger in one direction than the other (SAS Institute Inc. 1990). The value of skewness can be positive or negative and is unbounded. An ideal distribution has a skewness of zero. Kurtosis is a measure of the heaviness of the tails of a distribution. The value of Kurtosis must lie between -2 and $+\infty$ (SAS Institute Inc. 1990). Commonly, a maximum value of +2 is set to decide whether or not the model should be re-examined.

Even after the normal distribution of the residuals has been proved, the quality of the model cannot yet be assured. The model could still be predicting with different accuracy over a range of possible conditions. In order to detect possible trends, residuals should be plotted against predicted values, explanatory variables and other variables not in the model. If the model is unbiased, the residuals in these plots should be randomly scattered around zero. Formal statistical tests can be used to complement the visual inspection of residual plots. For example, a linear model of residuals on predicted values (or other explanatory variables) can be fitted. A significant slope parameter in such a linear model would indicate a significant trend of the residuals with predicted values or with the independent variable under evaluation. Zhao (1999), for instance, made extensive use of this approach for evaluating each model component. Alternatively, observed *versus* predicted values can be plotted and a simultaneous test for a slope of 1 and intercept of 0 can be performed (Mayer and Butler 1993; Gadow and Hui 1999).

For diameter distribution models or other models with tree- or diameter-class-resolution, specific evaluation procedures should be adopted (Reynolds *et al.* 1988). This topic is discussed in Section 7.7.4.

7.7 MODELS WITH TREE-LEVEL OR DIAMETER-CLASS-LEVEL OF OUTPUT RESOLUTION

As was mentioned in the Introduction (Section 6.1), three general approaches can be used to develop models with individual tree (or diameter class) resolution. These are (i) diameter distribution models; (ii) dis-aggregative models and (iii) individual tree models. These approaches are described in the following sections.

7.7.1 Diameter distribution models

Given a set of projected stand variables obtained from a stand model, diameter distribution models predict the number of trees per unit area by diameter class and the average height of each diameter class. Then, with an appropriate two-dimensional tree volume equation, per unit area yields can be estimated. For this purpose, it is often assumed that the underlying diameter distribution of a stand can be adequately characterised by a probability density function (p.d.f.). Several probability distributions have been used to model diameter distributions, including the gamma (Nelson 1964), log-normal (Bliss and Reinker 1964), Beta (Clutter and Bennett 1965; McGee and Della-Bianca 1967; Lenhart and Clutter 1971; Goodwin and Candy 1986), Weibull (Bailey and Dell 1973), and Johnson's S_b (Hafley and Schreuder 1977). However, most recent studies have used the Weibull distribution (Kuru *et al.* 1992). This function was first reported for forest diameter distribution modelling by Bailey and Dell (1973) and has been extensively used from then on (Clutter and Allison 1974; Schreuder *et al.* 1979; Matney and Sullivan 1982; Burk and Burkhart 1984; Woollons and Hayward 1985; Rodrigues da Cunha Neto *et al.* 1994).

The Weibull distribution is a 3-parameter function given by:

$$f(X)=\frac{c}{b}\left(\frac{X-a}{b}\right)^{c-1}\exp\left[-\left(\frac{X-a}{b}\right)^{c}\right] \qquad (7.7)$$

where a = location parameter,

b = scale parameter,

c = shape parameter, and

X = a random variable.

A closed form cumulative distribution (c.d.f.) can be readily derived:

$$F(X) = 1 - \exp\left[-\left(\frac{X-a}{b}\right)^{c}\right] \tag{7.8}$$

The location parameter *a* represents the minimum diameter of the stand. However, it should be noted that the tree of minimum diameter in the stand might not be encountered in the sampling process. As a result, the definition of the location parameter has varied widely depending upon the modellers. Some researchers have set the location parameter equal as the observed minimum diameter of the sample (Bailey and Dell 1973), or half of the observed diameter (Knoebel *et al.* 1986), or some proportion thereof (Frazier 1981). Projecting or modelling the minimum diameter is difficult because small trees in a stand are significantly influenced by genetic, microsite and silvicultural factors (Kuru 1989; Xu 1990).

Parameters *b* and *c* can be estimated by moments (García 1981), percentiles (Borders and Patterson 1990) or maximum likelihood (Smalley and Bailey 1974). Shiver (1988) compared the three methods and concluded that, under the likely assumption that there are no specific underlying parameters, both moments estimation and a percentile estimator reproduced the underlying distributions as well or better than maximum likelihood estimation.

Although in some early studies the parameters of the Weibull distribution were estimated by regression analysis on stand variables (Smalley and Bailey 1974), more recently modellers have preferred to recover the Weibull parameters from projected stand statistics (e.g. Hyink and Moser 1983; Knoebel *et al.* 1986). This approach, referred to as the parameter recovery approach, ensures compatibility between stand and diameter distribution estimates of basal area per unit area. Shortt and Burkhart (1996) compared the two methods and found that the parameter recovery model performed better for

shorter projection periods than the parameter prediction model, but the difference diminished with longer projection periods.

A refinement of the parameter estimation procedure of the Weibull distribution has been proposed by Lindsay *et al.* (1996). It involves the usage of skewness information in a three-parameter method-of-moments estimation technique, which avoids the problem of trying to calculate an appropriate proportion of the minimum diameter. A comparison of this procedure to the standard procedure of two-parameter method-of-moments estimation, revealed that the new approach was more accurate (Lindsay *et al.* 1996).

The Reverse Weibull, the location parameter for which is set as the maximum diameter (rather than as the minimum) has been proposed as a more logical way of representing diameter distributions (Kuru 1989; Xu 1990; Kuru *et al.* 1992; Mason 1992). The main advantage of the reverse Weibull as opposed to the usual Weibull is that the maximum diameter of the stand can be more readily predicted and is closely associated with changes in stocking (Kuru 1989). Additionally, locating the start of the diameter distributions at its maximum rather than its minimum results in a more accurate representation of the more valuable crop component, namely, the larger trees (Kuru *et al.* 1992).

Another advantage of using the reverse Weibull, that has been overlooked, is related to the simulation of thinning operations. While minimum diameter would be instantly changed with a thinning from below, maximum diameter would be unaffected. Therefore, with the reverse Weibull approach there would be no need to estimate changes in the extreme diameter at each thinning operation, thus avoiding a potentially large source of bias.

The 3-parameter reverse Weibull distribution is given by:

$$f(X) = (c/b)\left[\left(\frac{a-X}{b}\right)^{c-1}\right]\exp\left[\left(\frac{a-X}{b}\right)^{c}\right] \qquad (7.9)$$

with cumulative density function (c.d.f.):

$$F(X) = \exp\left[-\left(\frac{a-X}{b}\right)^c\right] \qquad \text{if } -\infty \le x \le a \qquad (7.10)$$

$$F(X) = 1 \qquad \text{if } x \ge a$$

The practical advantage of using the reverse Weibull over the conventional Weibull distribution was clearly demonstrated by Kuru (1989) and Xu (1990). The usage of skewness information in the reverse Weibull distribution could potentially improve the estimation of the maximum diameter (location parameter). This approach has not been explored yet.

There may be difficulties in detecting the extremes (maximum, minimum) of the population *dbh*. This is because only a tiny proportion of the population is sampled by each PSP. Therefore, the maximum diameter of a stand can be underestimated with the plot value and the minimum diameter in a stand can be overestimated (García 1991). The population variance estimated from the sample variance can also be biased (García 1991). Zhao (1999) addressed this topic and concluded that biases in estimations of stand diameter variances from plot variances were not serious for most stands but were serious for a few stands.

Intuitively, one would expect that with larger plots the magnitude of bias in estimating population extremes and variance of *dbh* would be reduced. Reynolds *et al.* (1988) simulated the effect of plot size on the precision of diameter distributions generated by the Weibull distribution using data generated with an existing growth model. As expected, the error index they proposed tended to decrease as plot size increased. Shiver (1988) also analysed the effect of plot size on the precision of diameter distributions using actual plot measurements and concluded that approximately 50 sample trees were needed to reproduce distributions into classes with less than 10% error in any class.

A solution to the problem of encountering the population maximum diameter (as opposed to the plot maximum diameter) has been proposed (Kuru 1989; Xu 1990). It is based on the extreme value theory and implies the use of another p.d.f. to represent the distribution of the maximum values. Xu (1990) and Xu *et al.* (1992) pointed out that the reverse Weibull itself is the appropriate distribution to be used for describing the distribution of

maximum diameters when the Reverse Weibull is being used for modelling diameter frequencies. Maximum diameters were stratified by tending regime (Kuru 1989) or age (Xu 1990) in order to compute the mean, maximum and standard deviation of the maximum diameters for each stratum. These three statistics were needed to solve the extreme-value-distribution parameters by the method of moments. From a methodological standpoint, this approach is very sound and sensible. For the implementation stage, however, it is not clear from the previous studies how the initial values of maximum-diameter statistics are to be obtained. When a plot is measured, the maximum diameter is readily available. However, the maximum value of the maximum-diameter-distribution is unknown unless a large number of similar plots have been measured. If such a large number of plots needed to be measured, then the absolute maximum of all plots in that particular stand could be used as the location parameter of the reverse Weibull. This latter approach was adopted by Zhao (1999) as the database at hand for that study contained a clear grouping of plots by stand at any given measurement age.

7.7.2 Stand table projection methods / dis-aggregative models

Despite the use of probability density functions (p.d.f.) being the more common approach to dis-aggregate stand statistics into diameter distributions, alternative approaches have been proposed. These are commonly referred to as *stand table projection methods*, *size class models* (Daniels and Burkhart 1988; Vanclay 1994), or *dis-aggregative models* (Ritchie and Hann 1997b). Their main advantage is that no predefined functional form (such as Weibull p.d.f.) has to be assumed. Instead, the initial tree list or stand table is projected forwards. Therefore, the real distribution of the diameters is taken into account. These models have characteristics of both individual-tree models and whole-stand models. The resolution of both input and output can be identical for dis-aggregative and individual-tree models, but functionally they are different. In dis-aggregative models, individual-tree information is aggregated as input to the whole-stand growth functions. The predicted stand growth is then dis-aggregated among trees in the sample list (Ritchie and Hann 1997b). This implies a tacit assumption that individual-tree information is irrelevant to the estimation of stand growth.

A number of dis-aggregation approaches have been proposed (Daniels and Burkhart 1988; Harrison and Daniels 1988; Nepal and Somers 1992; Zhang *et al.* 1993; Cao and Baldwin 1999). However, those based on the concept of relative basal area seem to be the ones that best combine good projection ability, biological foundation and mathematical simplicity. The relative basal area of a given tree is defined as the ratio of its basal area to the stand's mean tree basal area (Clutter and Allison 1974):

$$R_i = \frac{g_i}{\bar{g}} \Rightarrow R_i = \frac{g_i}{G/N} \tag{7.11}$$

The concept of relative basal area is very appealing because it provides a logical way to apportion stand-level estimates of basal area into diameter classes, ensuring complete compatibility. However, it has not been extensively used. The more well-known of the stand-table projection methods based on the concept of R_i is the one described by Pienaar and Harrison (1988) and Pienaar (1989). It is based on the expression proposed by Clutter and Allison (1974) to represent the trend of the relative tree size with time:

$$R_2 = R_1^{(T_2/T_1)^{\beta}} \tag{7.12}$$

where $R_i = g_i/\bar{g}_i$;

g_i is the basal area of the subject tree at time t_i (years),

$\bar{g}_i$ is the mean basal area per tree (G_i/N_i) at time t_i (years), and

β is a parameter to be estimated from individual-tree repeatedly measured data

When an estimate of β is available, a projected stand table consistent with the observed or projected total basal area (G_2) can be obtained. The procedure requires projection equations for number of trees per unit area (N_2) and basal area per unit area (G_2). The predicted mortality must be identified in the initial stand table. For this purpose, it is assumed that the probability that a given tree dies during the projection interval is inversely proportional to its relative size. This method performed better than the Weibull distribution parameter-recovery approach in the studies in which they were compared (Borders and Patterson 1990; Nepal and Somers 1992; Knowe *et al.* 1997).

Growth models using the concept of relative basal area were first reported by Clutter and Allison (1974) and then by Woollons and Hayward (1985). The latter study describes a dual simulator in which different levels of resolution (i.e. stand- and tree-level) are fully compatible. Stand level projection equations were used to estimate basal area, stand top height and number of stems per hectare. Tree-level models predict mortality proportions by diameter class and R_i. The selected R_i function was:

$$R_{2i} = \left(\frac{\overline{g_1^*}}{\overline{g_2}}\right)^{\alpha}\left[R_{1i}^* + \frac{\beta}{\overline{g_1^*}(T_2 - T_1)}\right] \tag{7.13}$$

where * indicates that the variable has been recalculated after allocating mortality

The parameters α and β in this formulation are plot-specific, whereas Clutter's β parameter applies for the whole population (or a large subset of it, e.g. thinned or unthinned stands). Another difference between Pienaar's stand table approach and the simulator described by Woollons and Hayward (1985) is that the former just assumes that the probability of mortality is inversely proportional to relative basal area, whereas the latter explicitly models the probability of tree mortality.

7.7.3 Individual-tree models

Forest growth models are often categorised into distance-dependent tree-level models, distance-independent tree-level models and stand-level models (Munro 1974). The individual tree models referred to in this section fall into the second category, distance-independent tree-level models. Additionally, more emphasis is placed on those which do not consider crown measures (e.g. crown ratio), as this variable is not always measured in common practice.

For developing a complete individual tree simulator, three basic components are required, namely (i) a diameter (or diameter increment) model, (ii) a model to predict tree

mortality/survival, and (iii) a model to predict individual tree heights. Each of these components is discussed separately in the following sections.

7.7.3.1 Tree diameter/diameter increment models

Two methodological issues are relevant for developing tree diameter models. One is concerned with the variable being modelled, i.e. diameter itself or tree basal area. The other involves a distinction between models that predict *increments* and models that directly predict the actual value at time t_2 of the variable being modelled (either diameter or tree basal area).

As pointed out by West (1980) two trees with the same diameter increment have different basal area increment if they differ in initial diameter. This fact might imply that the two variables express tree growth somewhat differently. West (1980) specifically analysed the practical implications of using diameter or tree basal area to model tree growth. He used two contrasting datasets (*Eucalyptus* species from Tasmania, Australia and three hardwoods from Ontario, Canada) and he consistently found virtually no differences in the results obtained by both methods.

Manley (1981) also compared diameter increment and basal area increment (along with their logarithms) in a study for *P. radiata* in New Zealand. He used Furnival's (1961) index for goodness of fit, as this statistic allows the comparison of models with different transformations of the independent variable, giving a measure of both the size of the residuals and possible departures from linearity, normality and homoscedasticity (Furnival 1961). Manley (1981) found that diameter increment yielded better values of Furnival's index and, thus, he chose that variable for his study (the logarithmic transformations had virtually the same values of Furnival's index as the actual increments).

Walsh (1986) examined a number of equation forms for predicting tree growth and concluded that, for that loblolly pine dataset, diameter increment was more appropriate than basal area increment. Hann and Larsen (1991) found that basal area growth

predictions were unreliable for very small trees, and used diameter increment as the response variable.

Diameter increment has been the preferred variable in New Zealand for describing individual tree growth of radiata pine in recent years (Gordon and Lawrence 1992, 1994; Shula 1997b). Elsewhere, other examples of using diameter increment include the studies by Knowe *et al.* (1997), Ritchie and Hann (1997a), Zhang *et al.* (1997a), Cao (2000) and Amateis *et al.* (1989). Tree basal area increment was used by Stage (1973), Murphy and Graney (1998), and Lynch *et al.* (1999).

Although most individual tree growth models predict increments rather than actual diameters (or tree basal areas) at time t_2, there are a few exceptions to this general rule. Zhang *et al.* (1996) used a model based on von Bertalanffy's differential equation claiming that it has certain desirable biological properties. Setting a number of assumptions allegedly concerned with factors affecting rates of assimilation (anabolism) and respiration (catabolism), Zhang *et al.* (1996) proposed the following 4-parameter projection equation:

$$dbh_2 = M\left\{1-\left[1-\left(\frac{dbh_1}{M}\right)^{(1/b_3)}\right]\exp\left[-b_1\ln(SI)(sp)^k\left(T_2^{b_2}-T_1^{b_2}\right)\right]\right\}^{b_3} \qquad (7.14)$$

where b_1, b_2, b_3 and k are regression parameters to be estimated;

M is the asymptotic maximum value of *dbh*;

sp is the average available space for each tree; and

SI is site index

Zhao (1999) fitted many traditional difference equations to *dbh* data. He also tried the following *dbh* projection equation derived from Clutter and Allison's (1974) relative basal area model:

$$d_2 = \sqrt{\frac{40000\bar{g}_2[R_1]^{(t_2/t_1)^\beta}}{\pi}} \qquad (7.15)$$

This equation requires stand basal area and stocking estimates at time t_2, and so it is not a truly tree-level model but rather a dis-aggregative way of allocating predicted stand basal area to individual trees. The contribution of projected stand statistics (G_2, N_2) in the individual-tree diameter growth equation, led to more precise projections than those obtained from sigmoidal difference equations of the type $d_2=f(d_1,t_1,t_2)$.

Masripatin (1998) compared diameter increment models and difference equations to model *dbh* of mixed tropical rain forests in Indonesia. She found that the Gompertz non-linear projection equation led to better results than linear diameter increment models, particularly over long intervals. In Masripatin's study, the projection intervals were constructed in a non-overlapping structure (see Section 7.5.4). According to suggestions of some stand-level modelling studies (Villanueva 1992; Lee 1998; Zhao 1999), the use a mixed-length interval structure would potentially improve longer projections.

Another potential advantage of using difference equations to model *dbh* is that the models so developed can be path-invariant. This is an important feature of any growth and yield model (Sullivan and Clutter 1972) that is usually neglected in individual tree models. Variations of the classic difference equations by including stand density indexes, site index and other site or stand attributes, might be an interesting alternative to diameter increment models which has not yet been extensively investigated for even-aged forests.

7.7.3.2 Tree mortality

Tree mortality has been categorised in regular competition-induced mortality and irregular or catastrophic mortality (Monsreud 1976). Catastrophic mortality might result from pests, diseases, fire, wind-throw, etc. and it is clearly more difficult to be modelled than competition-induced mortality.

Hamilton and Edwards (1976) suggested that the *logistic* function is the statistically preferred model for expressing the relationship between a dichotomous dependent variable (such as survival/mortality) and a set of independent variables. Monserud (1976) compared three methods (discriminant analysis, probit analysis and logistic analysis) and concluded that the logistic equation provided the greatest discriminating power for predicting live and dead trees. The logistic model has the following form:

$$P = \left(1 + e^{-X\beta}\right)^{-1} + \varepsilon \tag{7.16}$$

where P= probability of tree living or dying

X= vector of explanatory variables

β= vector of regression coefficients

ε= error term

The probability yielded by the logistic model is limited to the interval [0,1]. This general model has been used in nearly all recent individual-tree studies that include a tree-mortality component (Manley 1981; Lundgren and Gordon 1997; Shula 1997a; Murphy and Graney 1998; Lynch *et al.* 1999; Zhao 1999).

A problem that frequently arises in tree-mortality studies is the presence of unequal time periods between measurements. This problem has been approached in a variety of ways. Hamilton and Edwards (1976) divided the probability of death by the length of the growing period. As pointed out by Manley (1981) this approach is flawed as it neglects the principle of conditional probability. Monserud (1976) included the length of the growing period in the exponent of the logistic function as another explanatory variable. Manley (1981) assigned a "sampling factor" (or weight) to all live trees based on the total number of live trees in the plots multiplied by the length of their growing periods; i.e. the total number of "live tree years" divided by the total number of live tree sampled. For dead trees Manley (1981) assigned a weight of one as they can only die once. Lundgren and Gordon (1997) and Shula (1997b) used the following formulation to accommodate for different re-measurement intervals:

$$\text{P(survival)} = \left(\frac{1}{1 + e^{f()}}\right)^{i} \tag{7.17}$$

where *i* is the time interval between measurements

f() is the exponent of the logistic function

This formulation assumes that the probability of survival is constant with time which, as Lundgren and Gordon (1997) acknowledged, is not true. However, if the interval between measurements is not excessive (the authors excluded intervals longer than four years from their analysis) this assumption can be acceptable.

The implementation of tree mortality models can be deterministic or stochastic. A deterministic model implementation compares the probability of mortality (or survival) of each tree with a threshold value (e.g. 0.5). If the probability of survival is lower than this threshold then the tree is killed. On the other hand, in a stochastic model implementation, the probability of survival is compared to a uniformly distributed random variable confined to the interval [0,1]; if the probability of survival is greater than the random number, then the tree survives (otherwise the tree is killed). The result of this procedure will depend on the set of random numbers used in each simulation. Therefore, the models should be run a number of times and the average effect should be used. This procedure was used by Manley (1981).

The main advantage of the deterministic approach is precisely its deterministic nature that implies that the same result is obtained every time the simulator is run for a given plot. However, it has the drawback that it would kill all smaller trees and keep all bigger trees. In reality, though, trees of all sizes can potentially die as catastrophic or random mortality can occur along with the competition-induced mortality. A typical example is the effect of wind-throw (especially notorious in New Zealand) which is more likely to affect bigger trees. Although the stochastic approach overcame that problem, its random nature might be somewhat disconcerting for the user as different answers are obtained each time the model is run for a given plot.

An alternative approach that combines the main advantages of the deterministic and stochastic approaches is available. A tree in a plot represents a number of trees in a hectare. This number is known as the *expansion factor* and it results from the quotient between 10,000 and the plot size (expressed in square metres). The predicted probability of tree survival can be multiplied by the tree's expansion factor. With this approach, a tree's expansion factor is progressively reduced with each time-step through the simulator (Shula 1997a). This implementation approach makes use of the predicted probability of survival instead of arbitrarily assigning a status of dead or alive to each tree. For

instance, most of the smallest trees in the population, which are represented by (say) a single tree in a plot, with a low predicted probability of survival, is killed through this approach, but not all of them. Similarly, a minor proportion of the largest trees in the population is killed using this approach.

7.7.3.3 Individual tree heights

The nature of tree-height growth is less predictable than diameter growth. Additionally, height measurements usually include greater errors than diameter measurements, which means that data used for fitting individual tree heights usually include a proportion of negative height increments (Manley 1981; Gordon 1996). This suggests that the usage of height increment models is of limited use, which was found to be the case by Manley (1981) who obtained very poor results using this method.

One approach to the modelling of tree height utilises the well-known relationship between tree height and *dbh* that exists for a group of trees in a plot at a given time. A large number of models have been proposed to describe this relationship. For instance, Huang *et al.* (2000) reviewed the topic and listed 27 models. The following Petterson equation has been extensively used in New Zealand:

$$h = BH + \left(a + \frac{b}{dbh} \right)^{-2.5} \tag{7.18}$$

where h = tree height

BH = breast height (1.4 m in New Zealand; 1.3 m in Uruguay)

dbh = diameter at breast height

Note: Zhao (1999) found that an exponent of -5 (instead of -2.5) was better suited for radiata pine stands in the Canterbury region

Goulding and Shirley (1979) and Lawrence (1990) used this height-diameter relationship to predict individual tree heights. They predicted the parameter *a* from age and other stand attributes. Stand mean top diameter (MTD) can be estimated from predicted diameters and the corresponding mean top height (MTH) can be obtained from MTH-age curves. Having estimates of *a*, MTH and MTD the parameter *b* can be solved directly.

This approach ensures consistency between tree heights and MTH, that is, the predicted height of a tree with *dbh* equal to MTD would be exactly the same as the MTH.

Zhao (1999) developed a regional model for predicting individual tree heights from actual or predicted *dbh*'s using a similar approach. However, instead of solving the parameter *b* from the estimate of *a*, MTH and MTD, Zhao also predicted the *b* parameter from stand and site attributes. The variables he finally selected to predict parameters *a* and *b* of the Petterson equation were age, altitude, site index and stocking. As MTH was not explicitly included in this model, the consistency and compatibility of individual heights with the MTH was not ensured.

Staudhammer and LeMay (2000) used a similar approach to model height-*dbh* relationships for a number of species in British Columbia. In this study, the addition of a few stand variables (stocking, basal area and the ratio of *dbh* to maximum *dbh*) into the basic equations increased the accuracy of height predictions. Zhang *et al.* (1997b) selected dominant height, stocking and age as the explanatory variables for predicting the two parameters of their height-*dbh* model.

Individual heights have been also predicted by relating the tree-height to MTH ratio to (i) the tree's competition index and the average competition index of the 100 largest stems per hectare (Tennent 1981), or (ii) the ratio of the tree's diameter to the mean top diameter (Gordon 1996; Shula 1997c). Again, this approach ensures consistency between individual tree heights and MTH.

7.7.3.4 Individual tree models in New Zealand

Manley (1981) developed a distance-independent tree-level model for Kaingaroa Forest (Central North Island). This model generally outperformed a distance -dependent individual tree model developed by Tennent (1981), which demanded intensive measurement (including mapping of trees within a plot) for its development and implementation. Manley's model comprises a diameter-increment projection equation, a mortality function (logistic model) and a tree height estimation procedure. Manley's model also features a simple thinning algorithm, which makes the probability of a tree being selected for thinning inversely proportional to the cube of its diameter, that is

$p(t) = 1/d_i^3$. As with most individual-tree models, Manley's model projects diameter increments.

Zhao (1999) developed compatible tree-level stand-level models for radiata pine in Canterbury. The tree-level model consisted of a diameter growth equation and a mortality logistic equation. The diameter equation used by Zhao is a re-arrangement of the relative basal area projection equation proposed by Clutter and Allison (1974):

$$d_2 = \sqrt{\frac{40000\bar{g}_2[R_1]^{(t_2/t_1)^\beta}}{\pi}} \qquad (7.19)$$

where $\bar{g}_2$ is the average tree basal area (m^2/ha) of a stand at time t_2 and R_1 is the relative basal area of a tree as defined in Equation 7.11.

The average tree basal area at time t_2 ($\bar{g}_2$) is to be computed from stand basal area and stocking estimates, thus ensuring consistency and compatibility between stand- and tree-level outputs.

Individual-tree height projections were also developed in Zhao's study. With an appropriate two-dimensional tree volume equation, tree volumes were generated. Projections from stand- and tree-level models for a given plot were not necessarily the same. Therefore, under the assumption that plot-level projections were more accurate than tree-level ones, dis-aggregative adjustments to tree-level projections were required. As these adjustments were made directly on the outputs, they are much simpler than the adjustments proposed in other studies (Nepal and Somers 1992; Somers and Nepal 1994; Candy 1997; Cao and Baldwin 1999), in which the constraints are incorporated into the fitting procedure.

An individual tree model for *P. radiata* in all growth modelling regions of New Zealand has been recently completed (Gordon and Shula 1999). Diameter increment models are exponential equations of the form $\Delta D = \exp(x_0 + x_1V_1 ... + x_nV_n)$, where x_i are parameters and V_i are site-, stand- or tree-level variables (Shula 1997b). The probability

of tree survival was modelled by the logistic function (Shula 1997a). Tree heights were estimated by relating the tree-height/MTH ratio to the tree-diameter/MTD ratio. Indeed, this relationship was not assumed constant, but it was adjusted by a power function containing site, stand and tree variables (Gordon 1996; Shula 1997c).

7.7.4 Comparing approaches

In order to compare model performances at different levels of output resolution, different approaches can be followed. One possibility is to compare the models at stand- and tree-levels separately (Ritchie and Hann 1997a). If the stand-level projection ability of a (stand-level-based) diameter distribution model were compared to that of a tree-level-based aggregative model, one would expect that the former would outperform the latter. Conversely, when analysing the models' projection ability at the tree (or diameter-class) level, one would expect that the tree-level model would represent the shape of the diameter distribution better than the diameter distribution model.

Another possibility is to compare the projected diameter distributions against actual diameter distributions (Knowe *et al.* 1997). In this way, the model ability to project both, the shape of the distribution and the implied stand aggregates (i.e. basal area and stocking), are being compared at once. Formal statistical inference to test whether two distributions are samples of the same population can be conducted with statistical tests such as the Kolmogorov-Smirnov test or the Pearson chi-square test. For two diameter distributions to belong to the same population they must have similar numbers of trees in each diameter class. This implies that the two populations not only need to have a similarly-shaped diameter distributions but also similar basal area and trees per unit area.

A better approach is to compare the stand level projections as well as the simulated distributions. This was done by Manley (1981), who compared the model performance of two individual-tree models and a diameter distribution model[2]. In this study it was concluded that the Kolmogorov-Smirnov statistic was of limited use, reflecting only what was obvious from stocking and basal area comparisons.

[2] Only one of the individual-tree models was developed by Manley. The other two models (one tree model and one stand model) had been previously developed by other authors for approximately the same population.

Reynolds *et al.* (1988) revised thoroughly goodness-of-fit tests and model selection procedures for pdf-based diameter distribution models (but many of the concepts they addressed also apply to other modelling approaches). They pointed out many technical problems associated with goodness-of-fit tests and proposed a better way to compare and validate models yielding diameter-class outputs, a so-called *error index.*

This error index is calculated as the weighted sum of the absolute differences between predicted and observed numbers of trees in each diameter class. The weighting factor can be the tree volume or basal area of the average tree of each diameter class. Most of the traditional goodness-of-fit tests treat an error of one tree in the 30-cm class in the same manner as an error of one tree in the 10-cm class. This is a clear drawback of these tests, as the impact of errors occurring in bigger diameter class have a larger impact in the aggregate. By using a weighting factor in the error index proposed by Reynolds *et al.* (1988) this shortcoming is avoided.

7.7.5 Compatibility

Diameter distribution models derived by parameter recovery methods (Hyink and Moser 1983) ensure that explicit stand-level estimates are equal to their implicit counterparts derived by aggregating tree numbers and basal areas across diameter classes. Similarly, stand table projection approaches based on the concept of relative basal area (Clutter and Allison 1974; Woollons and Hayward 1985; Pienaar and Harrison 1988) also ensure that stand basal area and stocking obtained by aggregating tree-level outputs equal explicit (whole-stand-based) estimates of these variables.

Aggregation of tree numbers and basal areas from individual-tree-model outputs based on diameter increment models may result in stocking and basal area estimates that are different to estimations from stand-level models. If both types of models are available for the same species in the same region, this lack of compatibility may be misleading for the users. Having both types of models may be useful (to use one or the other according to the type of data available), but compatibility must be ensured to get consistent overall results regardless of the model chosen.

Compatibility between tree-level- and stand-level-based models may be achieved during the stage of fitting the model's parameters (Nepal and Somers 1992; Somers and Nepal 1994; Candy 1997; Cao and Baldwin 1999) or during model implementation (Zhao 1999). In both cases, a decision must be made as to which of the two models (i.e. tree-level-based or stand-level-based) is to be adjusted. Stand-level outputs will be available regardless of the basic modelling unit chosen. This, together with the fact that stand-level projections are generally considered more robust than tree-level projections, determines that tree-level estimates are the ones to be adjusted.

An example of dis-aggregative adjustments of tree-level estimates, taken from Zhao (1999) is displayed below:

- adjust n_{2i} so that $\sum n_{2i}^{*}=N_2$: $$n_{2i}^{*}=\left(N_2-\sum n_{2i}\right)\frac{n_{2i}dbh_{2i}^{-1}}{\sum n_{2i}dbh_{2i}^{-1}}$$

where n_{2i} is the unadjusted expansion factor (frequency) of the i^{th} tree;

n_{2i}^{*} are adjusted expansion factors; and

N_2 is the stand-level estimate of stocking.

- adjust d_{2i} so that $\sum g_{2i}^{*}n_{2i}^{*}=G_2$: $$dbh_{2i}^{*}=dbh_{2i}\sqrt{\frac{G_2}{0.00007854\sum dbh_{2i}^{2}n_{2i}^{*}}}$$

where g_{2i}^{*} are adjusted tree basal areas expressed in m^2; and

G_2 is the stand-level estimate of basal area (m^2/ha)

- adjust v_{2i} so that $\sum v_{2i}^{*}n_{2i}^{*}=V_2$: $$v_{2i}^{*}=v_{2i}\frac{V_2}{\sum v_{2i}n_{2i}}$$

where v_{2i}^{*} are adjusted tree volumes in m^3; and

V_2 is the stand-level estimate of volume (m^3/ha).

By weighting with the reciprocal of predicted diameter, frequency adjustments performed by Zhao (1999) were larger for smaller trees, which was consistent to the model of probability of tree survival (that predicted larger mortality for smaller trees). Tree diameters and volumes were uniformly adjusted based on differences between explicit and implicit estimates of stand basal area and volume, respectively.

Given that individual tree models typically need to be implemented through annual iterations, frequency and diameter adjustments must be made at each iteration.

7.8 BACKGROUND ON THE SPECIES STUDIED

7.8.1 *Eucalyptus grandis*

For its fast growth, good form and suitable wood for timber, pulp and fuel *E. grandis* is probably the most widely planted species of its genus for industrial wood production in the world, with an estimated area of plantations of about two million hectares in 1987 (Burgess 1988). This species has proved exceptionally successful in temperate and subtropical latitudes (FAO 1979; Franklin and Meskimen 1983; Garcia de Leon and Griffin 1995).

E. grandis is one of the main tree species commercially planted in Uruguay, with an area planted of 137192 ha as at 1999, representing 37.2% of eucalypt plantations, and 28.6% of all plantations (Echeverría 2000). Initially, most plantations of this species in Uruguay were aimed at pulp production, although in the last few years there has been a shift towards sawlog regimes.

In Uruguay, *E. grandis* has shown growth rates between 25 to 35 $m^3ha^{-1}an^{-1}$, although in some exceptional sites growth rates of up to 40 to 45 $m^3ha^{-1}an^{-1}$ have been recorded. Typical rotations range from 8 to 10 years for pulp (unthinned) regimes to 16 years for sawlog regimes. No growth and yield models are available for this species in the country, which results in thinning regimes and rotation ages being largely defined on an *ad hoc* basis.

7.8.2 *Pseudotsuga menziesii* (Douglas-fir)

Douglas-fir is one of the best known forest species in international markets, having an established reputation in the key world markets as one of the best timber species (Belton 2000). Its wood is very well suited to structural uses, given its low moisture content, good stability and high strength to weight ratio.

In New Zealand, Douglas-fir occupies the second place in terms of area planted after radiata pine, with 86 000 ha (5.0% of plantations) as at 1 April 1999 (Facts & Figures 2000). However, there has been a renewed interest in this species in the last few years, with more money being invested in establishing Douglas-fir in 1999 than in radiata pine (Belton 2000).

Douglas-fir plantations in New Zealand have been generally managed under longer rotations (typically 40 to 50 years) than radiata pine with later and commercial thinnings, and without pruning operations. In order to ensure acceptable quality levels for structural uses, branch suppression has been achieved by using high stockings. The high stockings that Douglas-fir stands can carry in New Zealand contribute to the extraordinary growth rates registered in the country, i.e. 16.5 $m^3ha^{-1}an^{-1}$ on average (Belton 2000).

The earliest growth and yield model for this species in New Zealand was developed by Mountfort (1978) for the Kaingaroa forest in the Central North Island (CNI). A whole-stand and a diameter distribution model for the same region were later developed by Xu (1990), using a much larger dataset. Following the same basic model design than for the radiata pine EARLY model (West *et al.* 1982, 1987), Fight *et al.* (1995[3]) developed DFEARLY, which has recently been upgraded (Leith Knowles 2001, pers. comm.). This model is largely driven by measures of crown length per hectare and is, therefore, particularly suited to early stages of stand development, when pruning and thinning operations are concentrated. Nonetheless, the model can be used to simulate stand growth over the entire rotation. The new version of DFEARLY can be calibrated for sub-regions within Kaingaroa (Leith Knowles 2001, pers. comm.).

[3] Unpublished paper presented to the 20th IUFRO World Congress, Tampere, Finland, August 6-12, 1995.

Douglas-fir models for the South Island of New Zealand include SIDFIR (Law 1990), DFIRSTAND (Temu 1992) and NEWDFIR (Lee 1998). No individual-tree models are currently available for this species in New Zealand.

7.8.3 *Pinus radiata* (radiata pine)

P. radiata is one of the most widely planted conifers world-wide, with most plantations located in the southern hemisphere (mainly in New Zealand, Chile and Australia). It is the main forest species in New Zealand with 1 563 000 ha planted (90.3% of plantations) as at 1 April 1999. Most of New Zealand's research on tree breeding, silviculture, growth modelling and wood processing has been centred on this species. In New Zealand, *P. radiata* achieves its largest growth rates in the world, with an average mean annual increment of around 24 $m^3ha^{-1}an^{-1}$.

Tending regimes for radiata pine in New Zealand have been distinctive, typically involving two waste thinnings and three pruning lifts up to 6 m, with final stockings ranging between 200 to 350 stems/ha over 28-year rotations.

Intensive measurements of permanent sample plots (PSP) monitoring tree growth on commercially managed plantations, as well as specially planned trials exploring extreme treatments, have been conducted over the last five decades. This has allowed development of extremely large databases that provide high-quality information for growth modelling purposes. This may explain the excellent international reputation of New Zealand in terms of growth and yield modelling expertise.

The earliest growth and yield model developed in New Zealand for *P. radiata* was developed by Beekhuis (1966). The second generation of models was based on the statistical analysis of PSP data measured up to 1973 (Goulding 1995). The KGM1 model was constructed for Kaingaroa Forest (CNI) and later modified to allow the use of regional site/top height/age curves to produce KGM2, a model that could be applied nationwide. The third generation of stand growth models was based on the state-space approach (García 1984, 1994) with an "interim" model for the CNI Pumice Plateau

(KGM3) being developed by Dunningham and Lawrence (1987). The latter model was later substantially upgraded and known now as the PPM88 model.

In addition to the previously mentioned models, which were developed at the former New Zealand Forest Research Institute (currently *Forest Research*), another model was independently developed by staff of NZ Forest Products Ltd. (NZFP). The NZFP growth model (Woollons and Hayward 1985) was compared with the KGM3 model by Hayward and Rawley (In: Hayward *et al.* 1987), who found that the KGM3 model overpredicted basal area growth of unthinned and lightly thinned stands in the Kinleith region. The NZFP growth model reported in Woollons and Hayward (1985) has been superseded by an unpublished upgraded version.

Submodels predicting growth responses to nitrogen fertilisation and level of genetic improvement (GF rating) have been developed as adjuncts to the PPM88 and other regional stand models (Goulding 1995).

A model that accounts for the effects of early thinning and pruning operations, largely driven by measures of crown length per hectare, have been developed by West *et al.* (1982, 1987). This model (EARLY) needs to be coupled with "later" stand models (as PPM88) after the stands receive their final silvicultural treatment or mean top height (MTH) reaches 18 m (whichever occurs first). EARLY simulations can be started from age 3 years or MTH of 3 m. For the CNI, simulations of hypothetical stands from age zero are possible by using the initial growth model (IGM) developed by Mason (1992) which has been recently updated (Mason 2001). This model simulates stand conditions (i.e. diameter and height distributions) up to age five years from information on altitude, initial stocking weed control, fertilisation, ripping, mounding and quality of seedling handling. Outputs from the IGM can be used as inputs of EARLY, which outputs can in turn be used as inputs of "later" stand models to simulate the growth and yield of hypothetical stands over an entire rotation. Alternatively, outputs from IGM can be inputted directly into stand models for older crops, without using the EARLY model (e.g. Mason *et al.* 1997). In this case, initial stands should be assumed to receive weed control in order to avoid severe underestimation of future yields (most PSP used to build models for mature crops receive adequate weed control).

Models such as KGM3 or PPM88 provide stand-level outputs, which can be disaggregated into diameter class information using the procedures described in Lawrence (1990). Individual tree models have also been developed in New Zealand (Tennent 1981; Manley 1981; Gordon and Shula 1999; Zhao 1999). Details of these models were given in Section 7.7.3.4.

CHAPTER 8

DATA DESCRIPTION

Data from three species were used in this study as mentioned in Chapter 6. The *Eucalyptus grandis* database was obtained from the network of trials of the National Forestry Programme[1] and from permanent-sample-plot data collected by four forestry companies in Uruguay.

The New Zealand databases (i.e. *Pinus radiata* and *Pseudotsuga menziesii*) were obtained from *Forest Research*'s permanent sample plot system (Dunlop 1995) after having sought and gained permission from the company owning the forests where the plots were located (Fletcher Challenge Forests).

8.1 DATA AVAILABLE - OVERVIEW

The *P. radiata* dataset was the largest of the three, and is possibly the largest for plantations worldwide. The Douglas-fir (*P. menziesii*) dataset consisted of fewer plots than the *E. grandis* one, but the former had many more measurements per plot than the latter. The Douglas-fir dataset had very good age coverage, whereas the *E. grandis* dataset was more confined to early ages. Of course, the average rotation ages for both

[1] A Division of the National Institute of Agricultural Research of Uruguay (INIA).

species are very different (10 to 16 years for *E. grandis* and 40 years or more for *P. menziesii*). A general description of the three datasets is shown in Table 8.1.

Table 8.1 General description of the three datasets

	E. grandis	*P. menziesii*	*P. radiata*
Number of PSP	423	251	973
Number of plot measurements	1125	2650	10520
Average meas. per plot	2.7	10.6	10.8
Min. meas. per plot	1	1	1
Max. meas per plot	9	20	27
Individual tree measurements	100367	189103	502817
Average number of trees/plot	79	103	75
Min. number of trees/plot	17	13	5
Max. number of trees/plot	300	327	613

The data coverage for the main variables is summarised in tabular form in Table 8.2. Several researchers recommended the use of graphical methods to describe the coverage of a dataset for growth modelling purposes (García 1984; Vanclay *et al.* 1995). The development of the mean top height (MTH) over time for the plots contained in the three datasets is depicted in Figure 8.1.

Table 8.2 Average, maximum and minimum values for the main variables.

	E. grandis			*P. menziesii*			*P. radiata*		
Variable	Mean	Max.	Min.	Mean	Max.	Min.	Mean	Max.	Min.
t (years)	5.2	15.8	0.9	33.8	84.2	7.3	18.0	72.0	0.4
MTH (m)	17.6	45.1	2.8	26.8	47.5	5.8	25.5	56.0	1.6
G (m^2/ha)	15.8	49.6	0.1	41.6	110.1	3.6	31.4	118.5	0.1
N (stems/ha)	850	1948	200	642	2850	44	479	8988	30
V (m^3/ha)	125	829	0	437	1480	14	321	1789	0
Mean *dbh* (cm)	15.6	35.5	1.5	32.9	69.8	7.5	32.2	91.3	1.1
Max *dbh* (cm)	21.5	47.6	3.3	47.6	96.1	15.7	43.8	116.8	2.6
Min *dbh* (cm)	9.1	28.5	0.4	18.0	49.5	0.1	18.8	75.9	0.0
Altitude (m)	-	-	-	450	786	195	468	976	50
Plot size (ha)	0.086	0.225	0.030	0.105	0.202	0.040	0.095	0.405	0.040

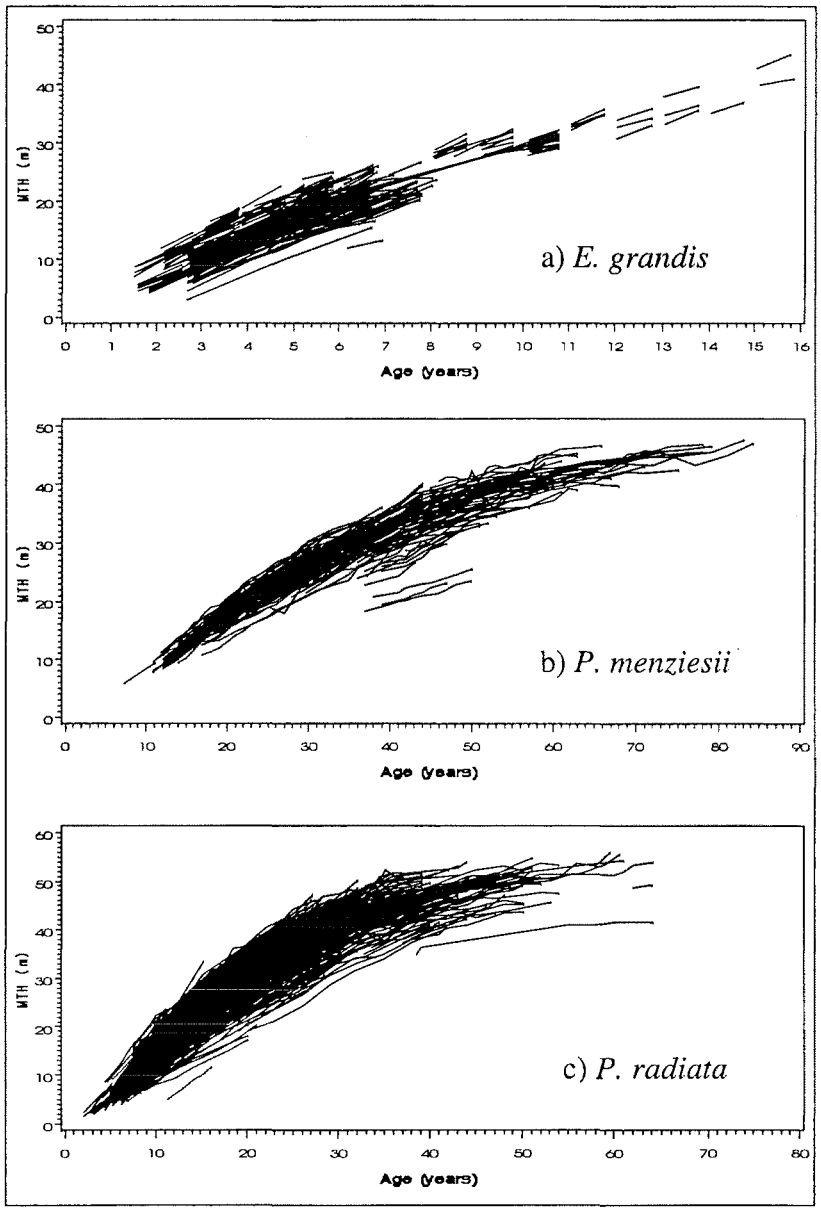


Figure 8.1 Mean top height development over time for the three datasets.

8.2 DISTRIBUTION OF PLOTS BY ZONE/FOREST

In Uruguay, there are four well-defined forestry regions, demarcated according to soil attributes, where the State promotes forestry plantations (Figure 8.2). *E. grandis* data come from three of these regions (i.e. Zones 7, 8 and 9), the most relevant for the plantations of this species. To put the dimension of these regions in perspective, the total land area of Uruguay is similar to the land area of the South Island of New Zealand. Uruguay has no mountains and climatic differences within the territory are minimal.

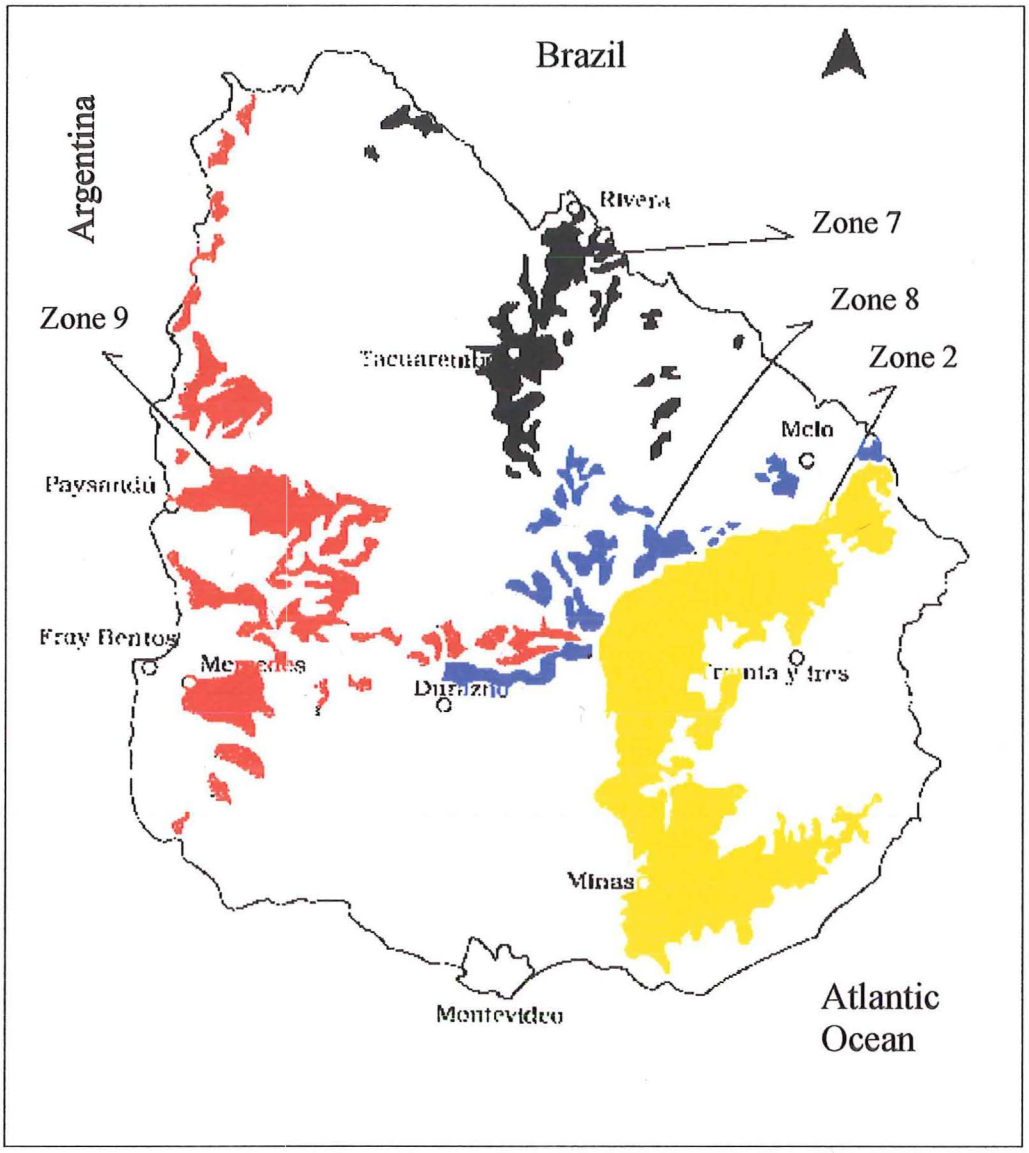


Figure 8.2 Location of forestry zones in Uruguay

Most plots were located in Zones 7 and 9 (Table 8.3), although there is still a reasonable number of plots from Zone 8 that may enable locality effects on growth patterns to be examined.

Both New Zealand datasets come from the Central North Island of New Zealand. This is a broad region where altitude effects on tree growth have been reported (Mountfort 1979; Mason 1992). There are a number of forests in the region, which may have particular patterns in terms of forest growth. Localising general growth trends by forest might be possible, although the fact that the majority of the plots come from the Kaingaroa Forest (76% for *P. radiata* and 86% for *P. menziesii*) virtually invalidates this approach (Table 8.3).

Table 8.3 Distribution of plots by zone or forest.

Species	Zone/Forest	PSP	Meas.
E. grandis	Zone 7	235	616
	Zone 8	28	109
	Zone 9	160	400
P. menziesii	KANG	217	2392
	WAKA	14	131
	WIRI	13	116
	HORO, ROEU, TAWE	2	6
	TAUH	1	2
P. radiata	KANG	759	9162
	TAWE	33	302
	ROEU	27	224
	WIRI	22	154
	MTHA	18	84
	WAKA, TAHO	15	186
	ROTO, TAUH	7	70
	MAMT	6	29
	NGAM	5	31
	TAHE, PINN, WAIO	4	75
	MANW, TUHO, STNL, WANI	3	47
	GAMM, SF28, OMAI, REWH, TUAR, NGAH, VELA, OMAN, CRTR	2	68
	HERE, WAIP, KERU, TANE, BROA, MANP, PAHU, RGTK, SF18, SUNV, TEWH, WIOK, WKWA, OPTR, PUTU, ROMA, TA8C	1	69

Note: for a description of the forest codes see Dunlop (1995), Appendix 9.

8.3 OVERALL DATA SCREENING

A general screening of the data was first performed for each dataset. This was carried out by extensive graphical plotting and by fitting simple models to each variable and inspecting extreme residuals. Suspicious measurements were singled out and checked. Some of them were corrected after identifying the cause of the error in original tree lists or field sheets. Others were tagged as suspicious and kept for further examination. Plot measurements showing catastrophic mortality due to cyclones were deleted.

Altitudinal effect on basal area and height growth in the CNI of New Zealand had been previously reported and was readily apparent when visually inspecting the graphs.

Therefore, those plots for which altitude information was missing were discarded. These represented only 5% (48) and 8% (20) of the plots for *P. radiata* and *P. menziesii*, respectively.

Plots with only one measurement could not be used for fitting most models (as these required interval-type data). However, they were appropriate for modelling height/dbh relationships as explained in Chapter 11.

8.4 CREATION OF INTERVAL DATASETS

For each species, all possible intervals were first created. Very short intervals (i.e. lower than 0.5 years) were eliminated as the impact of measurement errors could obscure actual growth for such short intervals.

8.4.1 Stand-level variables

Intervals with decreasing mean top height or basal area or increasing stocking were singled out for further checking. The equation form that best fitted each variable (e.g. Schumacher polymorphic for basal area) was fitted to each dataset and those observations having absolute residuals greater than 3.5 standard deviations were also singled out for further checking. For all suspicious measurements the original tree lists were checked. In some cases, the causes of outlier observations were punching errors or basal area or stocking calculations that were not taken from all trees in a plot, and these errors were corrected. However, the cause of a few suspicious measurements could not be established, even after consulting the manager of *Forest Research*'s PSP database (Mrs. Judy Hayes). As some plot measurements were dated 50 years back, checking the original records was not always feasible and a few suspicious measurements showing inconsistencies with other variables and/or with previous or posterior measurements were discarded. Only observations for which residuals from selected equations exceeded 4 standard deviations were deleted. In a normal distribution, only one observation out of 15779 would satisfy this condition (considering both tails). If retained, these extreme

outliers would produce undue effects on residual sum of squares and parameter estimates.

For the New Zealand datasets, intervals with a previous thinning in which thinning details were not available were removed. Only a minor proportion of the thinnings (1.6 and 5.2% for *P. menziesii* and *P. radiata*, respectively) did not include thinning details (i.e. basal area and stems thinned).

By virtue of having few measurements per plot in the *E. grandis* dataset, only 700 to 800 growth intervals, depending on the (stand-level) variable, could be formed. All of them were used for model fitting and a random sub-sample consisting of a single interval per plot was used for validating the significance of parameters.

In the case of the New Zealand datasets, the number of all-possible intervals was extremely high. In order to reduce the amount of autocorrelation and redundant data, a maximum of two (*P. radiata*) and four (*P. menziesii*) randomly selected intervals per plot were used for model development. Longer intervals were given increased probability of selection by multiplying each observation's random number by the factor *ln*(interval+2). Those two or four intervals (depending on the species) with largest values of that numerical product were selected. This weighting factor was selected as it generated the best distribution of intervals among a number of weighting factors explored. Datasets with little autocorrelation were formed with a single interval per plot (different to the ones used for model fitting), and were used to check the parameter significance of fitted models.

Only for *P. radiata* was it possible to set aside entire plots for validation purposes. The validation dataset was formed by 100 plots, 86 from the Kaingaroa forest and 16 from six other forests. For the other two species, validation was performed on intervals obtained from the same plots used for model fitting. In the case of *P. menziesii* intervals not used for model development were used for validation, whereas for *E. grandis* the same intervals were used for developing and validating the models. The type of data in the different datasets used for stand-level variables is summarised by species in Table 8.4.

Table 8.4 Summary of the type of data in stand-level datasets

	Species		
Dataset purpose	*E. grandis*	*P. menziesii*	*P. radiata*
Model development	all possible intervals	up to 4 interv. per plot	up to 2 intervals per plot
Checking param. significance	1 interv. per plot (included in the above dataset)	1 interval per plot (not included in the above dataset)	
Validation	all possible intervals (same plots used for model development)	intervals not used for model development (same plots)	all possible intervals from plots not used for model development

8.4.2 Tree-level variables

Tree-level intervals were used for modelling the growth of diameter at breast height (*dbh*), changes in relative basal area and probability of tree mortality. The approaches selected for modelling individual tree heights did not require the construction of growth intervals. As explained in Section 7.7.3.3, they were based on the relationship between tree height and *dbh*.

In general, diameter increment models are not path-invariant and need to be implemented through annual iterations. Therefore, emphasis should be placed upon maximising the accuracy and precision for annual predictions. The use of non-overlapping (short) intervals is thus to be preferred. On the other hand, models describing the trend of relative basal area and diameter difference equations can be path-invariant. For these components, the inclusion of longer intervals may provide better model capabilities, particularly over long projections (Lee 1998). Therefore, two types of tree-level intervals needed to be constructed.

The numbers of tree-level intervals were massive for all datasets, particularly for *P. radiata*. Extensive random sub-sampling was applied to reduce the number of observations to 5000 per species. For constructing mixed-length datasets, observations were sorted by interval-length classes. Different probabilities of selection were used to smooth the differences in number of observations between these classes.

For both non-overlapping and mixed-length intervals a single observation per tree was randomly sub-sampled, which reduced the level of autocorrelation in the data. Then, a maximum number of trees per plot (different for each species) was set to keep the number of observations close to 5000. This, in turn, avoided any risk of over-representing some plots in datasets used for model development.

Trees with decreasing diameters and observations consistently emerging as outliers after fitting various models were carefully examined. When punching errors were unambiguously identified (e.g. by checking the original field sheets for *E. grandis* in Uruguay) the observations were corrected. In a few cases, observations with unrealistic height/diameter relationships or clear inconsistencies with previous and posterior measurements were deleted. For no dataset were the deleted observations more than a dozen out of 5000 observations. Intervals containing dead trees at age t_2 were omitted for diameter and relative basal area modelling, but were kept for modelling the probability of tree mortality. As dead trees are recorded only once in *Forest Research*'s PSP system, observations for dead trees at further measurements needed to be created (for the construction of longer intervals).

Tree-level intervals not selected for model development were used for validating each model component. In the case of *P. radiata*, validation intervals were constructed only from the plots set aside for validation. Complete individual tree models (i.e. integrating all tree-level components) were also evaluated and compared against other modelling approaches using complete tree lists from selected plot intervals (Chapter 12).

8.5 DATA FOR MODELLING HEIGHT / DIAMETER RELATIONSHIPS

Height-*dbh* pairs with corresponding site and stand variables at a given time were needed for predicting tree height from *dbh*. Only actual height measurements were used to derive the relationship (not all trees are usually measured in each PSP). In order to minimise the level of autocorrelation, a single observation per tree was randomly chosen.

For a given tree in the *E. grandis* and *P. menziesii* datasets, height measurements not used for model development were used for validation. For *P. radiata*, tree height models were validated against height measurements obtained from the validation plots.

Methods and procedures used for this model component are described in Chapter 11.

CHAPTER 9

STAND-LEVEL AND DIAMETER DISTRIBUTION MODELS

9.1 INTRODUCTION

This chapter describes the construction of diameter distribution models, which are based on stand-level components for each of the datasets. The components required for a simple whole-stand model include (i) a mean top height equation, (ii) a net basal area/ha equation, (iii) a stocking equation (mortality) and (iv) a stand volume equation. Normally, the first three components would be projection (difference) equations, whereas the fourth component would by a static model based on some or all of the previous components.

To dis-aggregate stand-level estimates further into discrete diameter classes, a probability density function (pdf) can be used. Previous studies demonstrated the good performance of the reverse Weibull distribution for this purpose (see for example Whyte and Woollons 1992), and therefore, this was the only pdf considered here. The parameter recovery approach was preferred, as it ensures compatibility between stand-level estimates and the underlying distribution. Parameters were derived by method-of-moments because of its simplicity and good performance.

The classical method of moments relies on two moments of the Weibull distribution (i.e. mean, and standard deviation or variance) for deriving parameters b and c, once the

location parameter (a) has been estimated. Lindsay *et al.* (1996) proposed a three-moment procedure that allows estimating the three parameters of the Weibull distribution with no need to estimate the minimum diameter. The required third moment is the skewness of the distribution. The possibility of using this method with the reverse Weibull, not reported in the literature, was explored here.

In terms of model components, the use of the reverse Weibull requires projection equations for standard deviation (or variance) and maximum diameter. A skewness projection equation would substitute the maximum diameter equation in a three-moment approach.

9.2 METHODS AND PROCEDURES

As the method for estimating the parameters of the reverse Weibull distribution determines whether a projection equation for maximum diameter or skewness is required, this needed to be defined first. Reverse Weibull distributions with parameters obtained by both methods were fitted to actual plot measurements. Residual cumulative distribution functions (cdf), i.e. observed F(x) - predicted F(x), were plotted against predicted cdf's. By visually inspecting these plots it was possible to clearly determine that the best fit was achieved by setting the maximum diameter as the location parameter. Moreover, the skewness-approach showed three major shortcomings that endorsed the decision of not pursuing it any further. Firstly, the solution of a required an iterative algorithm. A Newthon-Raphson algorithm implemented through SAS's PROC MODEL was used here, but this further sophistication is undesirable for the implementation of the developed models. Secondly, the procedure yielded negative estimates of a in some cases. Lindsay *et al.* (1996) reported the same, but those authors could overcome the problem by setting a to zero in these cases (as they worked with minimum diameters). Thirdly, skewness trajectory over time was much more difficult to model than maximum diameter. The former was sometimes decreasing and sometimes increasing, whereas the latter was (of course) consistently increasing and easy to model.

In the case of *P. menziesii*, the effect of *Phaeocryptopus* and how it should be accommodated given the limited amount of data from the pre-disease period needed to be explored.

For the three datasets, a number of difference equations (both anamorphic and polymorphic) derived by the algebraic difference approach from classical growth models were fitted to each variable (all equations are listed in Appendix 1). Sigmoid equations were fitted to mean top height, basal area, maximum diameter and standard deviation of diameters. Inverse sigmoid equations were fitted to stocking.

For each variable, a few models showing the best fit (lowest MSE and mean residual, and unbiased residual plots against predicted values and independent variables) were further refined and compared. Improvements were generally made through linear modifications of the original parameters, following the principles described in Section (7.6.7). For stand mortality, different data structures along with Woollons' 2-step approach (see Section 7.6.3) were also performed.

For developing stand volume models (Section 9.6) "observed" under-bark volumes needed to be calculated. At each PSP measurement, observed volumes were obtained by summing up individual tree volumes, which were in turn calculated with appropriate tree volume equations. The tree volume equation for *E. grandis* was derived from the modified Max and Burkhart (1973) taper model (Equation 2.11a with parameters in Table 5.2). Appropriate volume tables (equations) for each plot measurement of the New Zealand datasets had already been assigned in *Forest Research*'s PSP system (Dunlop 1995). Volumes for all measurements of *P. menziesii* were calculated with the Tree Volume Table 136. Various volume tables were used for the *P. radiata* dataset, although Table 10 was the most common. Details of these equations can be found in the software FFCALC Version 1.3 (Forest Research Institute 1992), while the equations and their parameters are displayed later as Equations 12.1 and 12.2.

9.3 EFFECT OF *PHAEOCRYPTOPUS* (*P. MENZIESII*)

The Xu (1990) model for *P. menziesii* in the Central North Island (CNI) of New Zealand (DFCNIGM3) caters for the effect of *Phaeocryptopus* on the growth and yield of the species. This was done by simply stratifying the model depending on whether the stand was assumed infected or not, which in turn was defined according to the year of measurement. Measurements prior to 1963 (year in which the disease started to spread in the CNI region) were assumed unaffected by the disease, whilst measurements after that year were assumed affected.

Basal area growth was found more affected than height growth (Xu 1990). Considering that and the fact that basal area is usually the key component to predicting future yields (Mason 1992), this variable was selected to analyse the effect of *Phaeocryptopus*. Only non-overlapping intervals (without any subsampling) were used for this analysis in order to detect more accurately the year at which basal area growth started to decline.

Among all difference equations tried, a two-parameter Schumacher polymorphic (Clutter 1963) fitted the data best. Residual plots clearly showed that residuals for measurements prior to 1966 were mostly[1] positive (basal area underestimated) and of larger absolute magnitude than residuals for later years (Figure 9.1).

Considering that only 5.1% of plot measurements (134 out of 2650) were recorded before 1966 it appeared extremely difficult to model the "pre-disease" period properly. Moreover, it can be reasoned that the developed model will always be applied to diseased stands, as this fungus is already established in the Central North Island and its complete eradication is unlikely. Therefore, it was decided to discard all measurements taken prior to 1966 and fit all components to data collected after that year.

[1] A few negative residuals for measurements prior to 1966 come from plots with high mortality, for which basal area growth was overestimated.

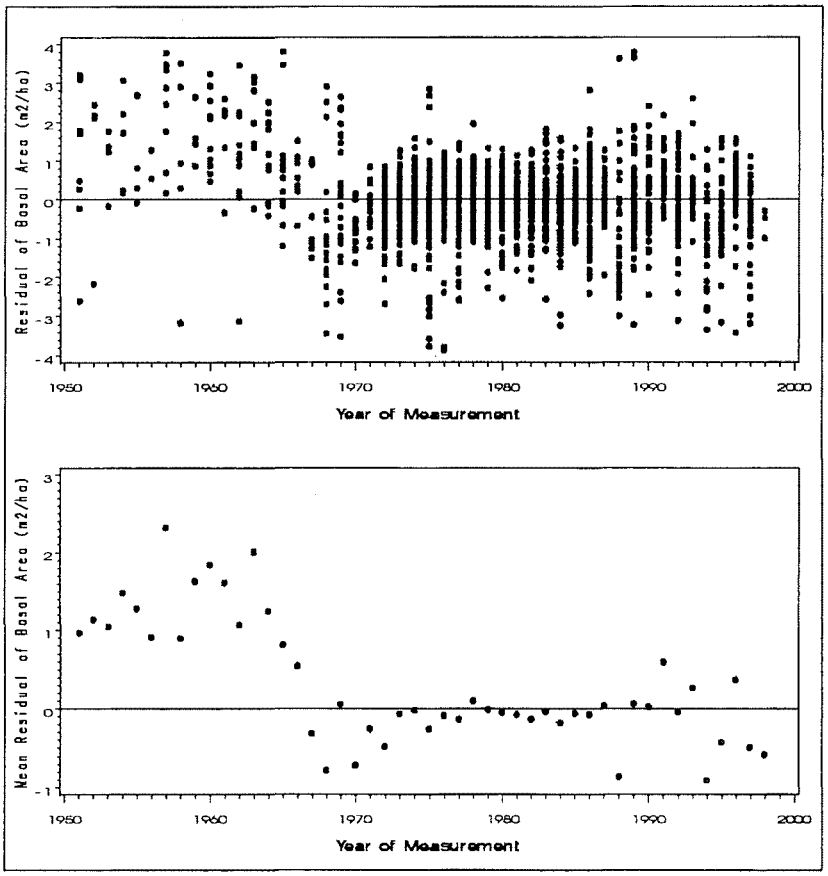


Figure 9.1 Basal area residuals (top) and mean residuals (bottom) against year of measurement.

Simply for the purpose of quantifying the growth loss caused by *Phaeocryptopus*, stand net basal area was modelled using a dummy variable that accounted for the presence/absence of the disease (this equation is presented in Appendix 2). The disease effect on mean top height and stocking could not be properly modelled. Projections for diseased and disease-free stands were made with the adjusted basal area equation and MTH, stocking and stand volume equations fitted to post-1966 data. Assuming starting values of 16 m^2/ha of basal area, 13 m of MTH, 800 stems/ha at age 15 and altitude of 400 m, projections to age 40 with and without the disease yielded stand volumes of 514 and 591 m^3/ha, respectively. These figures represent mean annual volume increments of 12.8 and 14.8 m^3/ha/yr respectively, with a yield reduction caused by *Phaeocryptopus* of approximately 13%. This reduction is slightly lower than yield reductions reported by Xu (1990).

9.4 MODELLING SIGMOID-TYPE VARIABLES

9.4.1 Mean Top Height (MTH)

The same polymorphic version of the Chapman-Richards model fitted the data best for the three datasets. This equation was obtained by letting the shape parameter of the Chapman-Richards be the free parameter, as explained in Section 7.4.3.2. The resulting equation is:

$$H_2 = a\left(\frac{H_1}{a}\right)^{\frac{\ln(1-\exp(-kt_2))}{\ln(1-\exp(-kt_1))}} \qquad (9.1)$$

For *P. menziesii* and *P. radiata*, parameters a and k were made linear functions of altitude, which removed the trends of the residuals with that variable and reduced the Mean Square Error (MSE) by 4.2% and 23.2% respectively. The effect of altitude was more apparent for *P. radiata* (Figure 9.2). Starting values at age 15 years used in Figure 9.2 are average values in the datasets for selected altitudes. Decreasing height growth of *P. radiata* with increasing altitude was also reported in previous studies (Mason 1992; Woollons *et al.* 1998) and would be largely explained by the lower temperatures at higher altitudes. For *E. grandis*, parameters a and k were modified by dummy variables that catered for differences in growth trajectories between zones (Figure 9.3).

Parameter estimates for the three datasets are presented in Table 9.1. All parameters were significant ($p<0.05$) when checked with minimum-autocorrelation datasets (see Section 8.4.1).

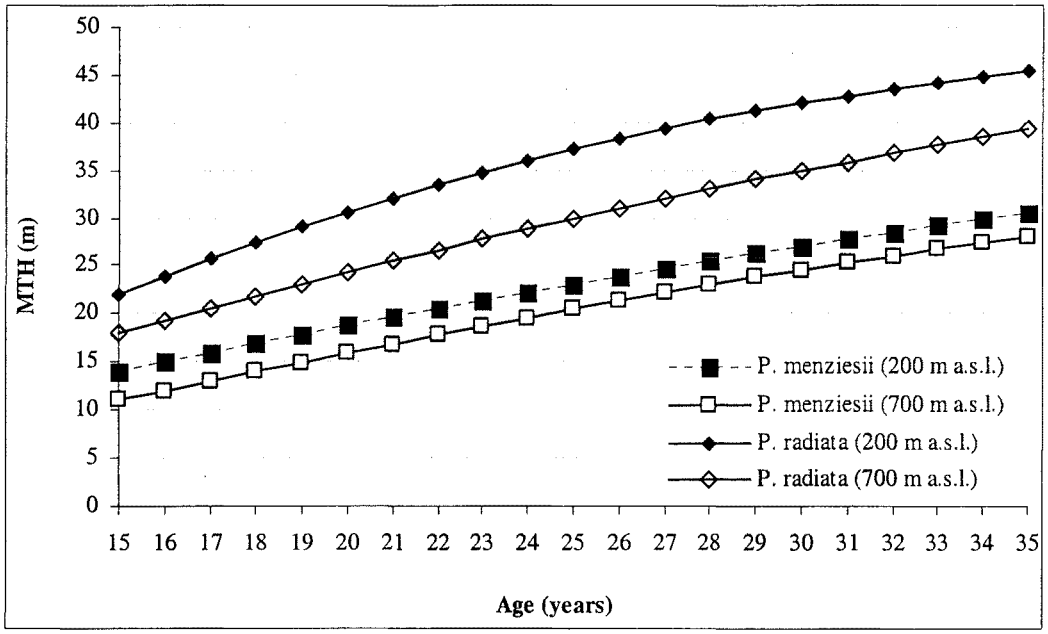


Figure 9.2: Altitude effect on MTH growth of *P. radiata* and *P. menziesii*.

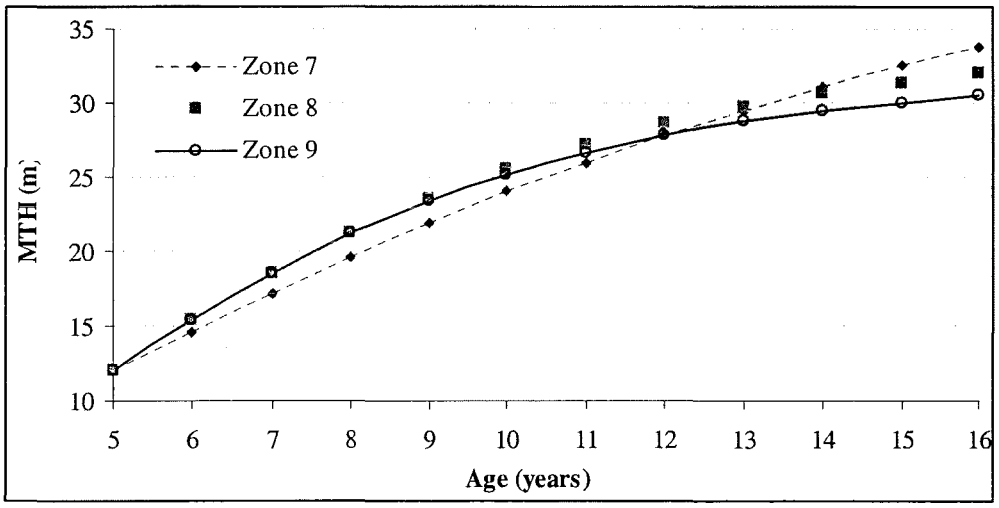


Figure 9.3 Projected MTH growth trajectories by zone (*E. grandis*).

Residual statistics and plots for the fitting subsets of the three species are shown in Table 9.2 and Figures 9.4 to 9.6, respectively. For *E. grandis* and *P. menziesii* residuals were labelled by zone and forest, respectively. Residuals were not labelled for *P. radiata*, as there were 25 forests in the dataset and no particular pattern was observed for any forest. Most observations (84% in the radiata pine dataset) were from Kaingaroa forest.

Table 9.1 Parameter estimates and standard errors for MTH models.

Parameter	*E. grandis*		*P. menziesii*		*P. radiata*	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
a	$a=a_0+a_1D_1+a_2D_2$		$a=a_0+a_1(ALT/100)$		$a=a_0+a_1(ALT/100)$	
k	$k=k_0+k_1D_1+k_2D_2$		$k=k_0+k_1(ALT/100)$		$k=k_0+k_1(ALT/100)$	
a_0	47.4374	2.1953	61.5105	1.509	48.1076	0.735
a_1	-12.8577	5.285	-2.2049	0.2652	1.9624	0.1724
a_2	-15.3666	2.4824	-	-	-	-
k_0	0.0983	0.00807	0.0219	0.00174	0.1034	0.00277
k_1	0.1243	0.0556	0.00283	0.000373	-0.00935	0.000502
k_2	0.1581	0.0204	-	-	-	-

Note: D_1=1 for Zone 8, 0 otherwise. D_2=1 for Zone 9, 0 otherwise

Table 9.2 Statistics of the residuals for MTH models.

Residual statistic	*E. grandis*	*P. menziesii*	*P. radiata*
Mean	0.13	-0.06	-0.03
Standard deviation	1.07	1.11	1.38
Skewness	-0.10	-0.14	-0.11
Minimum (1%)	-2.40	-3.11	-3.77
Maximum (99%)	2.55	2.75	3.42

Note: Minimum (1%) and maximum (99%) are the first and 99th percentiles of the residual distribution, respectively. They are not absolute minima and maxima.

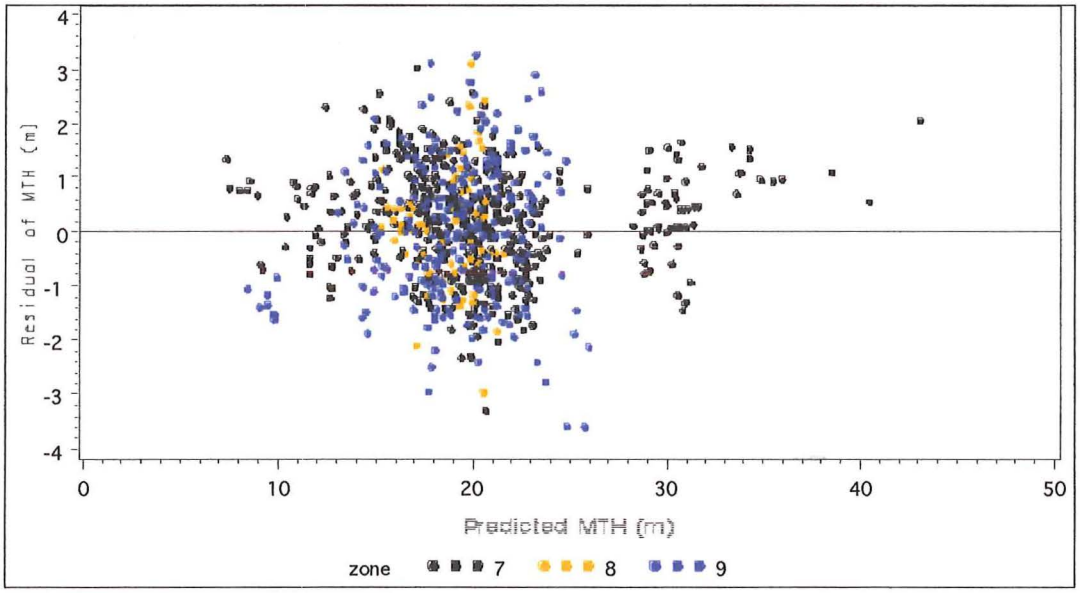


Figure 9.4 MTH residuals labelled by zone (*E. grandis*)

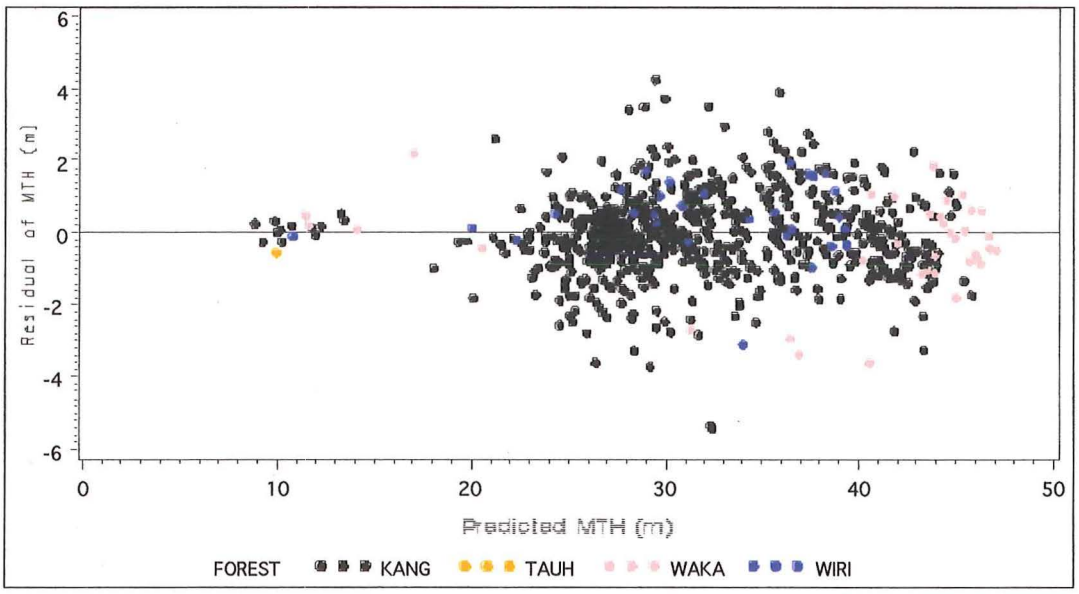


Figure 9.5 MTH residuals labelled by forest (*P. menziesii*)

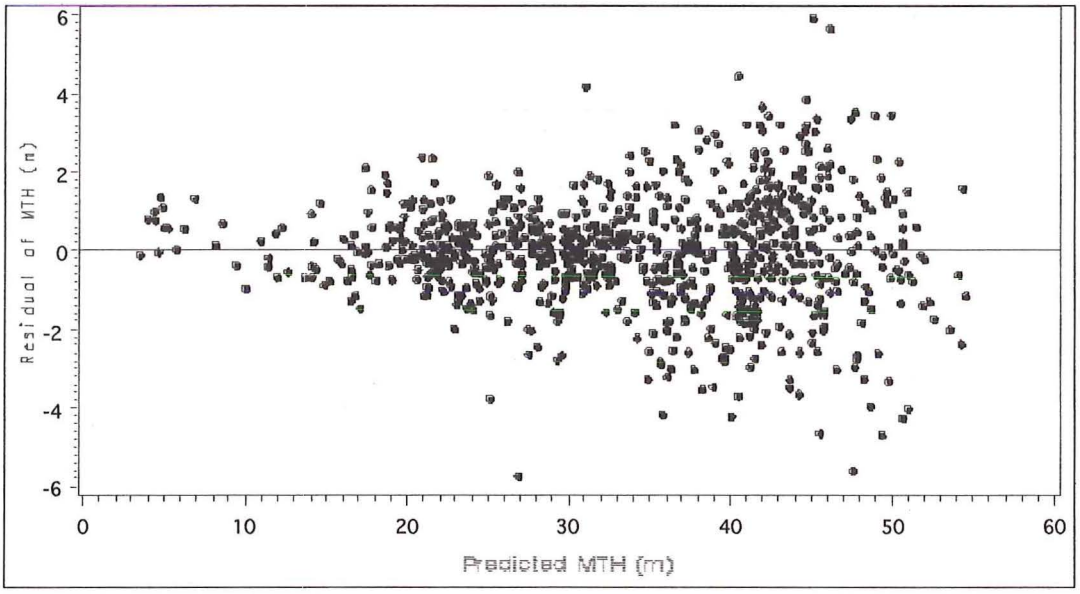


Figure 9.6 MTH residuals (*P. radiata*).

9.4.2 Net Basal Area (G)

As in the case of MTH, the same model form was found best for the three datasets. This was the following 2-parameter Schumacher polymorphic equation:

$$G_2 = G_1^{(t_1/t_2)^b} \exp\left\{a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\} \qquad (9.2)$$

This Schumacher equation was also found to be a good fit to several basal area datasets in a comparison of five sigmoid equations conducted by Woollons and Wood (1992). For the one-parameter version of this model (i.e. without *b*) Clutter (1963) made the asymptote parameter (*a*) a linear function of site index. This was tried here but the site index parameter was not significantly different to zero ($p<0.05$). However, altitude was successfully introduced in the model as a linear modification of the asymptote parameter for *P. radiata* and of both parameters for *P. menziesii*. Dummy variables for locality were used for *E. grandis*.

For *P. menziesii*, the effect of thinning on basal area growth was modelled, using the ratio of the proportion of basal area removed at thinning to thinning age (i.e. (Gt/Gb)/tt). The parameter associated with the thinning term was positive (Table 9.3), indicating that if there are two stands with the same basal area at a given age, one just thinned and the other unthinned, then the thinned stand will grow faster than its unthinned counterpart. This is contrary to the accepted interpretation that the loss of crown surface through thinning would reduce basal area growth (West *et al.* 1982; García 1988a). *P. menziesii* stands in New Zealand are typically managed with high stockings and late thinnings (relative to *P. radiata* stands). This determines that there can be significant proportions of sub-dominant and suppressed trees, which are more affected by *Phaeocryptopus* than dominant trees. As all thinnings were thinnings from below, the removal of suppressed trees may pre-empt tree mortality, which constitutes a direct reduction of net basal area, and improve the overall stand vigour. Temu (1992) also found a positive effect of thinning on basal area growth for *P. menziesii* in New Zealand. By having the age of thinning (tt) in the denominator of the thinning term, the effect will be lower for late thinnings. Thinning details for the thinning term correspond to the last thinning, which implies that the asymptote will tend to return to its original value for successive thinnings. In order to avoid unrealistic projections from very early and intense thinnings, the thinning term could be limited to its maximum value in the dataset (i.e. [Gt/Gb]/tt = 0.0593).

Table 9.3 Parameter estimates and standard errors for basal area models

Parameter	*E. grandis*		*P. menziesii*		*P. radiata*	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
a	$a=a_0+a_1D_1+a_2D_2$		$a=a_0+a_1\frac{ALT}{1000}+a_2\frac{Gt/Gb}{tt}$		$a=a_0+a_1*(ALT/1000)$	
b	$b=b_0+b_1D_1+b_2D_2$		$b=b_0+b_1(ALT/1000)$		$b=b_0+b_1\sum(Gt_i/Gb_i)/tt_i^2$	
a_0	4.3769	0.0643	3.6362	0.0352	4.6599	0.0340
a_1	-0.7519	0.0810	3.0225	0.1162	0.2225	0.0547
a_2	-0.5461	0.0832	5.1912	0.5729	-	-
b_0	0.8027	0.0315	1.8378	0.0550	1.1402	0.0241
b_1	1.2005	0.1227	-2.0074	0.0837	-5.4665	0.6242
b_2	0.4350	0.0675	-	-	-	-

Note: thinning terms are defined in the text. D_1 and D_2 as defined in Table 9.1

For *P. menziesii*, the inclusion of altitude and the thinning term individually reduced the MSE by 33.7% and 12.1%, respectively. When the two factors were combined the reduction in MSE totalled 41.3%, indicating that the improvements caused by each factor were almost completely retained. There seems to be an interaction of the two effects, whereby the positive thinning response is more evident at lower altitudes (Figure 9.7). A possible reason for that would be the higher incidence of *Phaeocryptopus* at lower altitudes (promoted by the warmer temperatures) which would also explain the higher levels of mortality observed in unthinned stands at low altitudes (see Section 9.5).

In contrast, basal area growth was reduced by thinning for *P. radiata*. For the latter species, the selected thinning term modified the rate parameter (b) and was computed by the following summation:

$$\sum \frac{(Gt_i/Gb_i)}{tt_i^2}$$

where Gt_i is the thinned basal area at the i^{th} thinning; Gb_i is the basal area before the i^{th} thinning; and tt_i is the age of the i^{th} thinning. i = 1, 2, 3 (a maximum of three thinnings were available in the dataset).

More thinnings could be simulated safely as the thinning term would only vary marginally with thinnings at late ages. The maximum value of the summation in the dataset was 0.2047

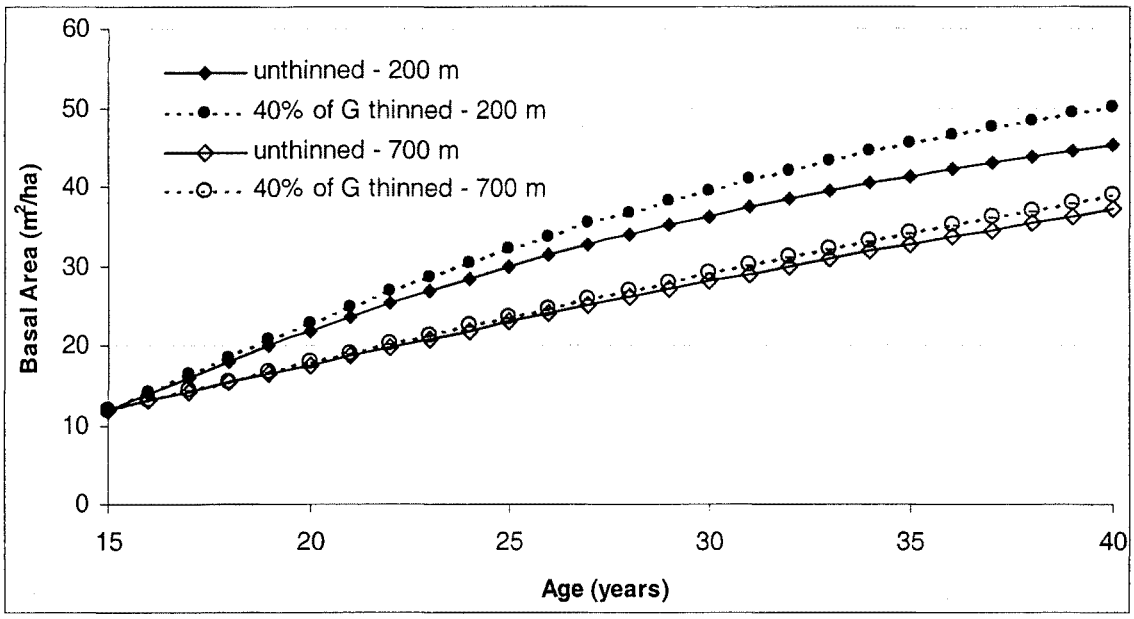


Figure 9.7 Projected basal area growth for unthinned stands and stands thinned to 40% of basal area at age 15 for two altitudes (*P. menziesii*).

This thinning term reduced the MSE of the model by 7.5% individually, and the reduction totalled 9.0% when altitude was also included (altitude alone reduced the MSE by 3.6%). In order to analyse the combined effects of altitude and thinning, basal areas at age 5 were simulated with the Initial Growth Model (IGM), Version 2 for *P. radiata* in the Central North Island (Mason 2001) for two contrasting altitudes (200 and 600 m a.s.l), initial stocking of 1000 stems/ha and weed control. Predicted basal areas at age 5 were projected up to age 30 without thinning and with a thinning reducing basal area by 50 % (Figure 9.8).

Initial basal area curves obtained from IGM were smoothly joined with basal area curves for unthinned stands between 5 and 30 years obtained using the basal area model developed here. Basal area/ha estimates at age 30 for unthinned stands were 18 and 22% larger than basal areas for thinned plots, for altitudes of 200 and 600 m respectively.

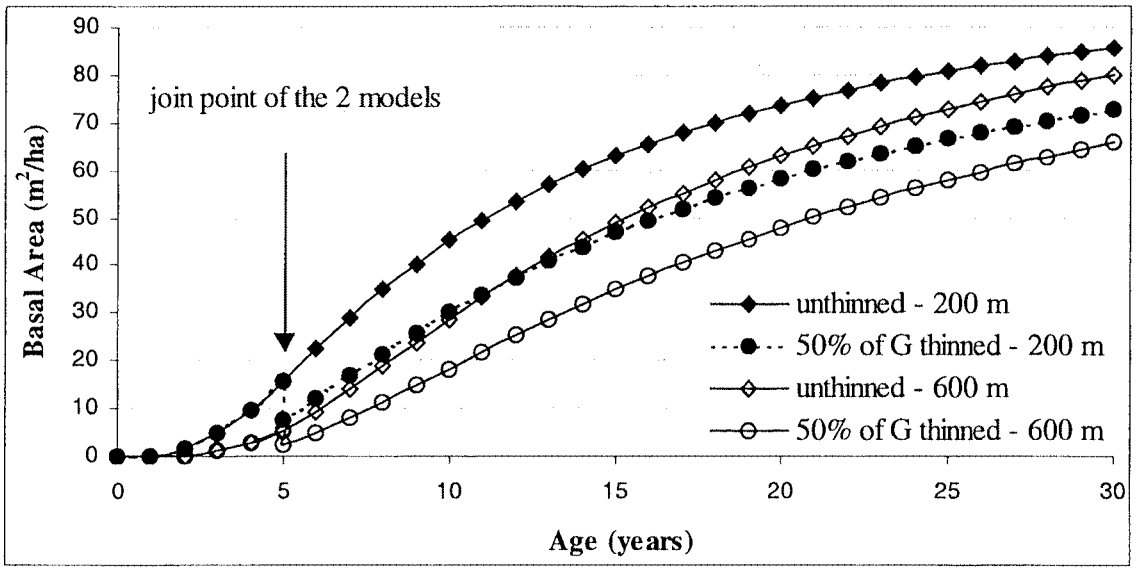


Figure 9.8 Projected basal area growth for unthinned stands and stands thinned to 50% of basal area at age 5 in two altitudes (*P. radiata*).

A thinning index describing the kind and intensity of thinning, proposed by Bailey and Ware (1983), was the most appropriate to characterise the thinning effect on basal area growth of *E. grandis*. The index modified the asymptotic parameter downwards, indicating a negative effect of thinning on basal area growth (as for *P. radiata*). Thinned plots in the *E. grandis* dataset were scarce and were only measured up to a few years after thinnings, which precluded ascertaining long-term effects of thinning. Locality adjustments performed with dummy variables had much greater impact on the MSE than the thinning term. Individually, the inclusion of dummy variables and the thinning term reduced the MSE by 42.6% and 19.6% respectively. The reduction achieved by the two effects combined totalled 49.5% suggesting that a great deal of the two effects was confounded, which was the case as most thinned plots were from Zone 7. For these reasons it was decided not to keep the thinning term in the final basal area model for this species.

Statistics and plots of basal area residuals for the fitting subsets of the three species are shown in Table 9.4 and Figures 9.9 to 9.11, respectively.

Table 9.4 Statistics of the residuals for basal area models.

Residual statistic	*E. grandis*	*P. menziesii*	*P. radiata*
Mean	-0.13	-0.08	0.11
Standard deviation	1.36	3.41	4.86
Skewness	0.16	0.20	-0.28
Minimum (1%)	-3.61	-8.26	-12.98
Maximum (99%)	3.68	9.25	12.13

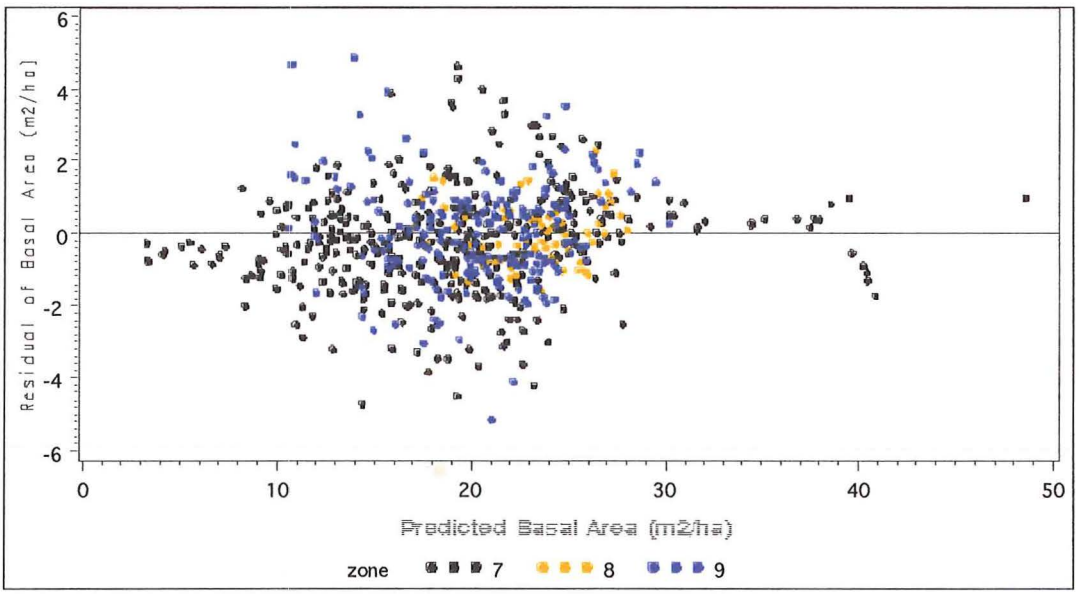


Figure 9.9 Basal area residuals labelled by zone (*E. grandis*)

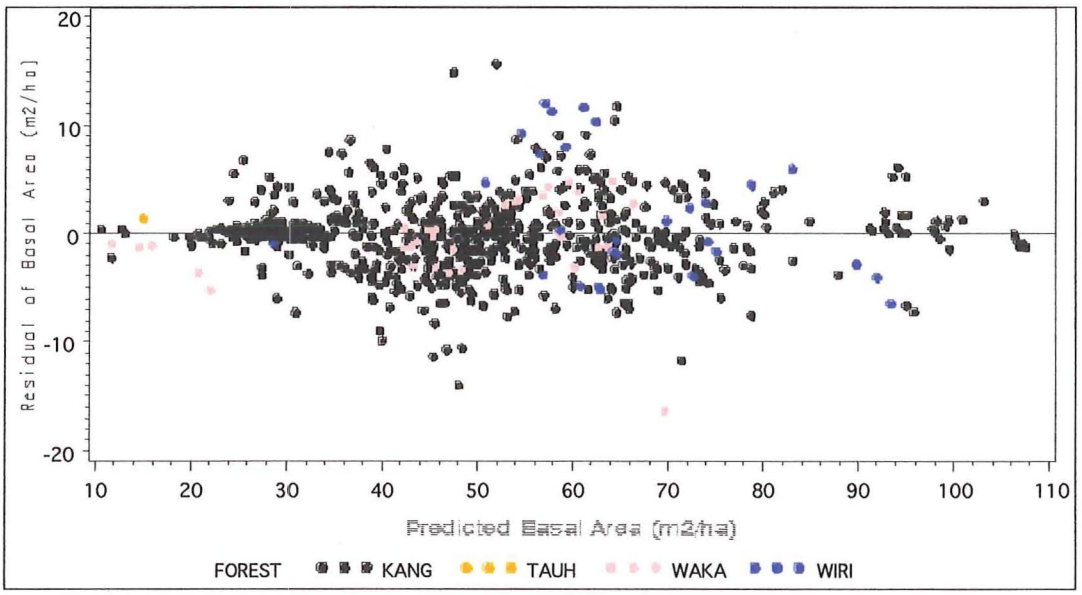


Figure 9.10 Basal area residuals labelled by forest (*P. menziesii*)

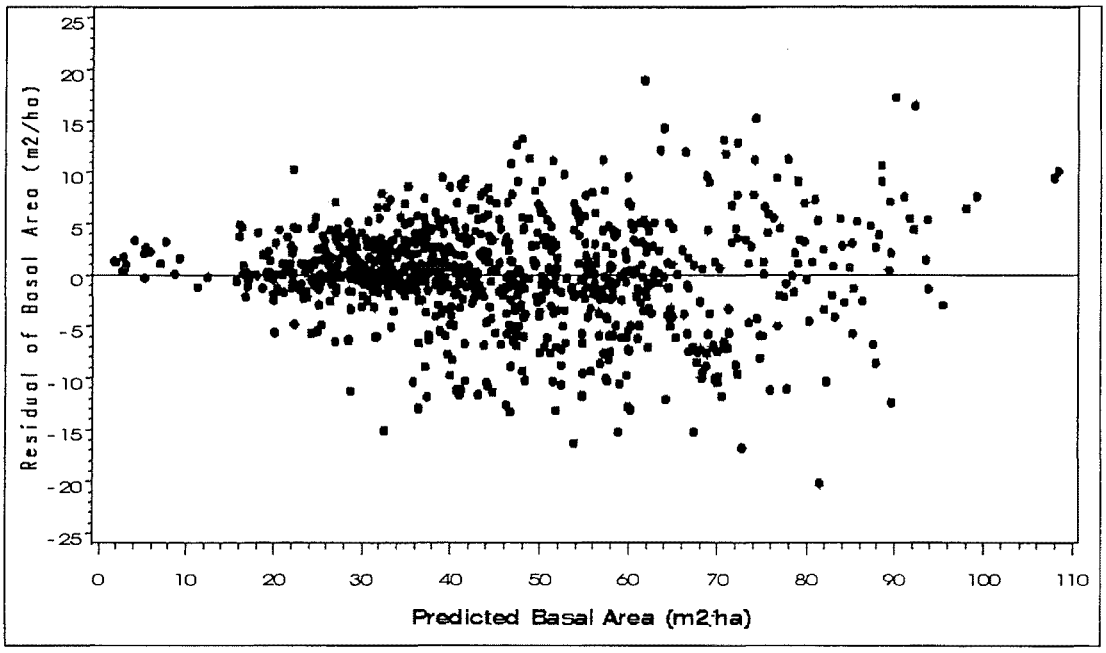


Figure 9.11 Basal area residuals (*P. radiata*)

9.4.3 Maximum Diameter (D_{max})

Maximum diameter was best modelled with individually different forms of model for the three species. The selected models included variations of the monomolecular equation (*P. menziesii*), Schumacher equation (*P. radiata*) and the same polymorphic version of the Chapman-Richards equation that was selected for modelling MTH (*E. grandis*).

Maximum diameter was closely and inversely related to stocking. However, the inclusion of stocking at any age T_i in the models would make them lose the path-invariance property, as stocking is not a constant. Instead of stocking, a related variable (*nn*) was successfully introduced in the models, removing the trends of residuals with stocking. This variable is the stocking at plot establishment or after thinning, that is, the starting[2] stocking unaffected by mortality.

The selected models for maximum diameter are displayed below with corresponding parameter estimates presented in Table 9.5.

[2] The term *starting* stocking is used here to avoid confusion with *initial* stocking, which usually refers to the number of planted trees per hectare.

E. grandis: $$D_{max2}=a\left(\frac{D_{max1}}{a}\right)^{\frac{\ln(1-\exp(-bt_2))}{\ln(1-\exp(-bt_1))}} \quad (9.3)$$

where $a=a_0+a_1*(\text{Site Index})+a_2/(nn/1000)$

P. menziesii: $$D_{max2}=D_{max1}\exp[-b(t_2-t_1)]+a\{1-\exp[-b(t_2-t_1)]\} \quad (9.4)$$

where $a=a_0+a_1*\text{altitude}$

$$b=b_0+b_1*(\text{altitude}/1000)+b_2/\ln(nn)$$

P. radiata: $$D_{max2}=D_{max1}^{(t_1/t_2)^b}\exp\left\{a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\} \quad (9.5)$$

where $a=a_0+a_1/nn$;

Table 9.5 Parameter estimates and standard errors for maximum diameter models.

	Parameter estimates (standard errors)					
Species	a_0	a_1	a_2	b, b_0	b_1	b_2
E. grandis	9.5673	0.7478	4.2543	0.2344	-	-
	(2.436)	(0.0885)	(0.1941)	(0.00938)	-	-
P. menziesii	132.866	-0.0585	-	-0.00988	0.0217	0.1287
	(5.1885)	(0.00767)	-	(0.00136)	(0.00197)	(0.0112)
P. radiata	5.0441	60.5201	-	0.5554	-	-
	(0.0356)	(3.1608)	-	(0.0145)	-	-

Residual statistics are shown in Table 9.6, and plots of residuals against predicted values and interval length (years) are displayed in Figures 9.12, 9.13 and 9.14. No severe trends are readily apparent, confirming the good fit of the selected models, even for very long projection periods.

Table 9.6 Statistics of the residuals for maximum diameter models (fit subsets).

Residual statistics	*E. grandis*	*P. menziesii*	*P. radiata*
n	705	800	935
Mean	-0.01	-0.21	0.46
Standard deviation	0.9	2.52	3.70
Skewness	0.40	0.25	-0.17
Minimum (1%)	-2.0	-6.0	-9.7
Maximum (99%)	2.4	7.0	10.0

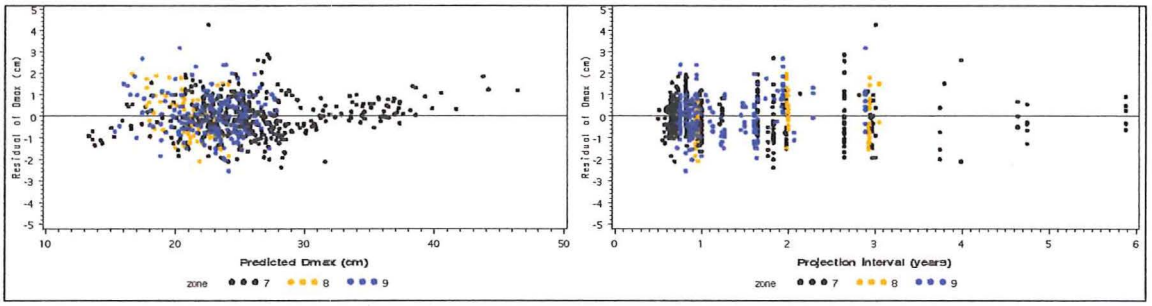


Figure 9.12 Residuals of maximum diameter against predicted values (left) and interval length in years (right). *E. grandis*.

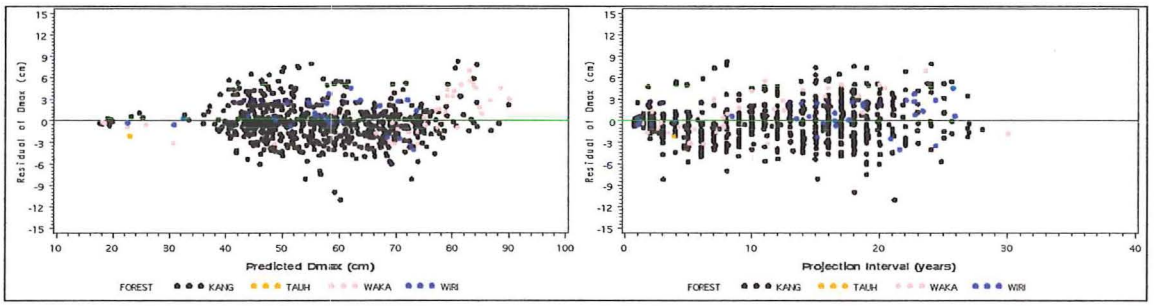


Figure 9.13 Residuals of maximum diameter against predicted values (left) and interval length in years (right). *P. menziesii*.

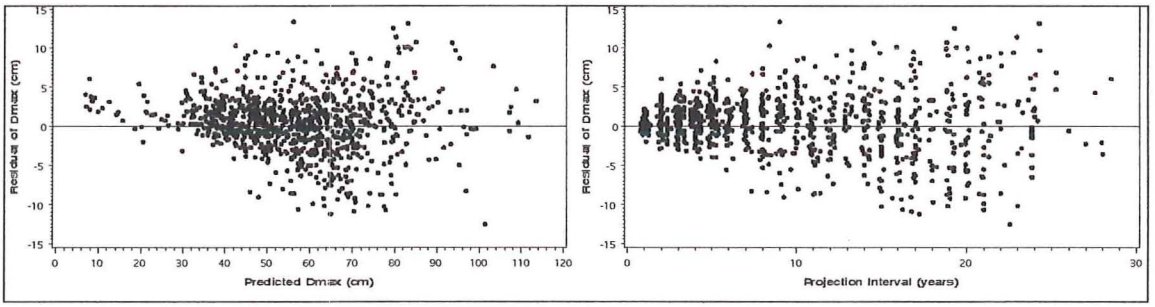


Figure 9.14 Residuals of maximum diameter against predicted values (left) and interval length in years (right). *P. radiata*.

9.4.4 Standard Deviation of Diameters (D_{std})

The same polymorphic version of the Chapman-Richards model selected for modelling MTH was also preferred for modelling the standard deviation of diameters for the New Zealand datasets. A 3-parameter polymorphic Gompertz equation was selected for *E. grandis*, although the Chapman-Richards model also fitted the data very well, ranking second among all model forms tried. The final models were:

E. grandis:

$$\begin{aligned} Dstd_2 &= \exp\left\{\ln(Dstd_1)\exp\left[-b(t_2-t_1)+c\left(t_2^2-t_1^2\right)\right]\right\} \\ &\exp\left\{a\left[1-\exp\left(-b(t_2-t_1)+c\left(t_2^2-t_1^2\right)\right)\right]\right\} \end{aligned} \tag{9.6}$$

P. menziesii and *P. radiata:* $$Dstd_2 = a\left(\frac{Dstd_1}{a}\right)^{\frac{\ln(1-\exp(-bt_2))}{\ln(1-\exp(-bt_1))}} \tag{9.7}$$

where $a=a_0+a_1*\ln(nn)+a_2*\ln(nn)^2$, for *P. menziesii*, and

$a=a_0+a_1*nn/1000$

$b=b_0+b_1*nn/1000$, for *P. radiata*

For the New Zealand datasets, it was found that the standard deviation of diameters decreased with increasing stocking. This relationship was weak but the incorporation of starting stocking (*nn*) clearly improved the fit of the models. In fact, MSE was reduced by 7.7 and 9.4% when this variable was added to the *P. radiata* and *P. menziesii* models, respectively. In the case of *P. menziesii*, the stocking effect in the model was best reflected through a second-degree-polynomial on the logarithm of *nn*. This explains why the parameter estimates were so different to those of *P. radiata* (Table 9.7). For *E. grandis*, no variable could be added to improve the fit of the Gompertz model.

Table 9.7 Parameter estimates and standard errors for standard deviation models.

	Parameter estimates (standard errors)				
Species	a, a_0	a_1	a_2	b_0, b	b_1, c
E. grandis	1.9356	-	-	0.2464	0.00714
	(0.0608)	-	-	(0.0149)	(0.000977)
P. menziesii	-74.9099	27.2668	-2.0597	0.0302	-
	(2.5008)	(0.9589)	(0.0823)	(0.0017)	-
P. radiata	15.2649	8.985	-	0.0302	-0.00421
	(0.6707)	(1.3716)	-	(0.00193)	(0.00174)

Residual statistics of standard deviation models are shown in Table 9.8, and plots of residuals against predicted values and interval length (years) are displayed in Figures 9.15, 9.16 and 9.17. Overall, the models displayed no trends with predicted values or interval length.

Table 9.8 Statistics of the residuals for standard deviation models (fit subsets).

Residual statistics	*E. grandis*	*P. menziesii*	*P. radiata*
n	706	797	942
Mean	0.02	-0.05	0.05
Standard deviation	0.27	0.84	1.22
Skewness	-0.08	-0.38	-0.38
Minimum (1%)	-0.8	-2.7	-3.4
Maximum (99%)	0.8	2.5	3.1

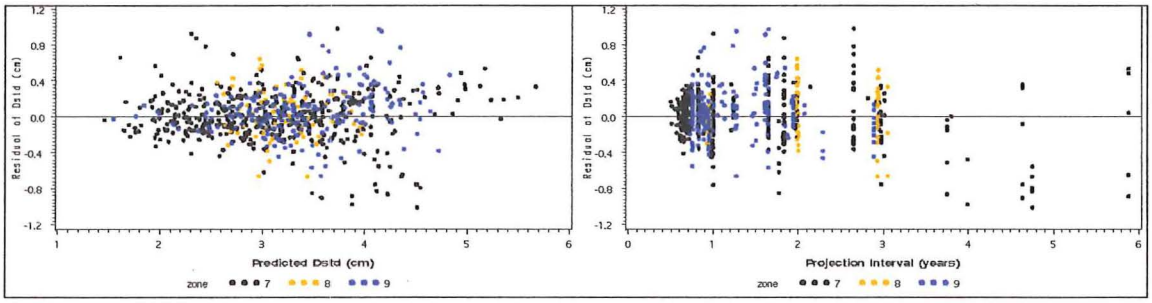


Figure 9.15 Residuals of standard deviation against predicted values (left) and interval length in years (right). *E. grandis*.

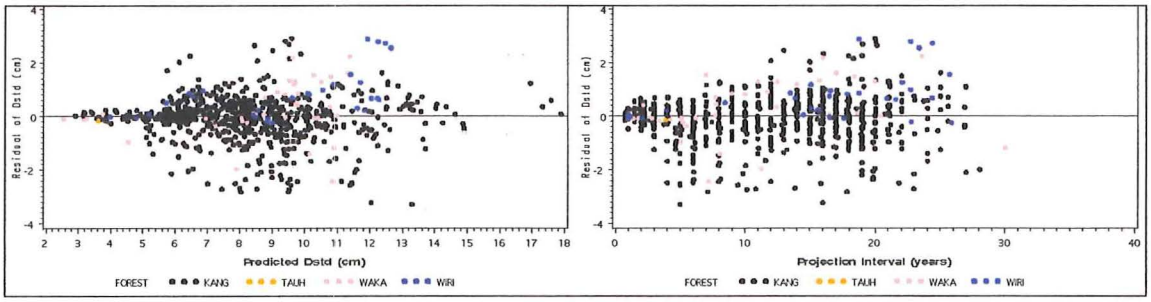


Figure 9.16 Residuals of standard deviation against predicted values (left) and interval length in years (right). *P. menziesii.*

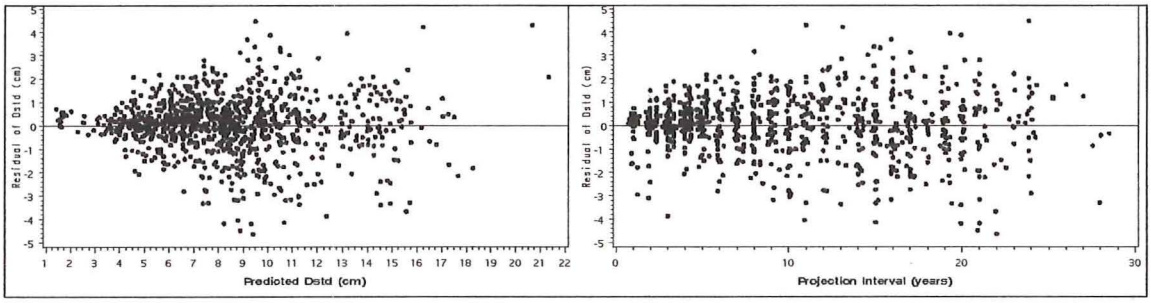


Figure 9.17 Residuals of standard deviation against predicted values (left) and interval length in years (right). *P. radiata.*

9.5 Modelling Stand Stocking (N)

A number of inverse sigmoid functions were compared for modelling changes in stand live stocking over time. This variable is very difficult to model, mainly because (i) not all intervals exhibit mortality, and (ii) because of the unpredictable nature of some factors (i.e. factors other than normal competition-induced mortality such as windthrow, pests, etc.) causing mortality.

Given this complexity, a further comparison of alternative modelling approaches was considered relevant. The three approaches compared were: (i) fitting the models to projection intervals randomly selected, regardless of whether they included mortality; (ii) Woollons' (1998) 2-step approach (see section 7.6.3); and (iii) fitting the models to the longest interval available for each plot.

The third approach was conceived as a fair compromise between using all selected intervals and using only intervals that exhibited mortality. By selecting only the longest

interval for each plot the probability of having some mortality is increased. However, no serious bias is being introduced as the same criterion is being applied to all plots. Moreover, given the path-invariance and consistency properties of the selected models, the predicted mortality would be monotonically decreasing for shorter intervals.

As explained in section 7.6.3, the two-step approach requires a logistic model to predict the probability of tree death. To fit this model, Woollons (1998) used data with constant interval length (one year) to avoid bias. However, the interval length was explicitly included here as an independent variable in the logistic regression. This allowed predicting the probability of tree death occurring for any period. As an example, probabilities of stand mortality occurring for *P. menziesii* are displayed in Figure 9.18.

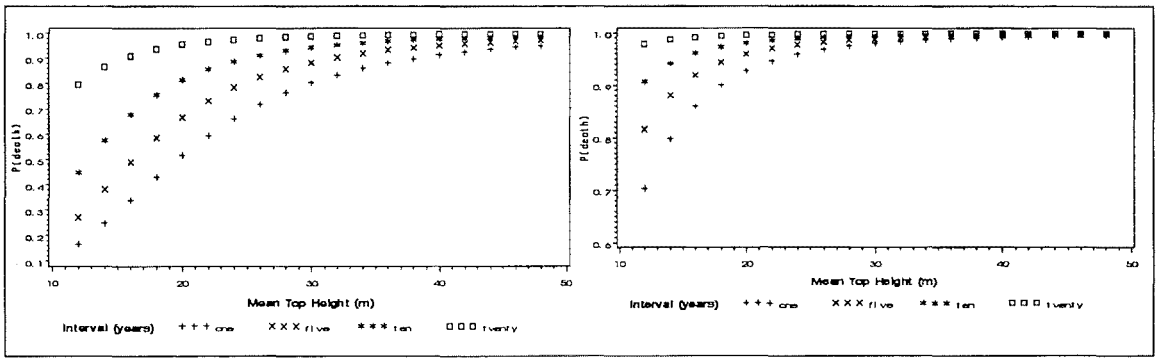


Figure 9.18 Probability of stand mortality occurring against MTH for 800 stems/ha (left) and 1600 stems/ha (right). Note the different scale for the y-axis.

Models developed through the three approaches were applied to the validation subsets[3] and residuals were scrutinised by graphical and numerical means. Residual statistics for the three species are shown in Table 9.9.

Table 9.9 Residual statistics for the three approaches to modelling stand stocking (validation subsets).

	E. grandis			*P. menziesii*			*P. radiata*		
Residual statistics	all interv.	2-step	longest	all interv.	2-step	longest	all interv.	2-step	longest
n	706	706	706	12448	12448	12448	5675	5675	5675
Mean	-1.3	-5.3	-0.6	0.9	-6.3	-0.3	4.9	-1.0	5.5
Standard deviation	31.5	31.9	31.5	37.2	39.2	38.4	26.2	25.4	24.6
Skewness	-0.7	-0.6	-0.7	-0.8	-1.8	-1.4	-1.8	-2.5	-0.6
Min (1%)	-89	-93	-87	-133	-157	-146	-112	-125	-85
Max (99%)	64	70	64	-266	110	114	67	55	75

[3] For *E. grandis* the validation subset is actually the same subset used for model fitting (see Section 8.4.1)

The relative merit of each approach varied with the dataset. For *P. menziesii* and *E. grandis*, the 2-step approach was clearly biased, overpredicting overall mortality (higher negative mean residual) and showing lower precision (higher standard deviation of residuals) than the other methods (Table 9.9).

For *P. radiata*, on the other hand, the 2-step approach generated the lowest overall bias and highest precision (Table 9.9). However, this approach presented the maximum amplitude between the first and 99th percentiles of the residual distribution indicating a more asymmetric distribution of residuals (confirmed by the highest absolute value of the skewness statistic). This increase in the asymmetry caused by the 2-step approach was also evident for *P. menziesii*.

Models fitted to the longest interval produced the lowest mean residuals for the validation datasets of *P. menziesii* and *E. grandis*. The performance of this method was also acceptable for *P. radiata*, despite the mean residual being greater than when the model was fitted to all selected intervals. In fact, precision (standard deviation), and distribution of residuals (skewness, balance between first and 99th percentiles) for the *P. radiata* validation dataset were best with the longest-interval approach (Table 9.9). Therefore, this approach was adopted for all three datasets. Selected models with parameter estimates and standard errors are displayed in Table 9.10.

Table 9.10 Selected stocking models

		Parameter estimates (standard errors)			
Species	Model	α	β	γ	δ
E. grandis (Eq. 9.8)	$N_2 = N_1\left(\frac{t_2}{t_1}\right)^{\beta} \exp\left[\alpha(t_2 - t_1)\right]$	-0.0352 (0.00466)	0.0449 (0.0203)	- -	- -
P. menziesii (Eq. 9.9)	$N_2 = \left[N_1^{\theta} + \frac{\beta + \delta W}{ALT * 1000000}\left(t_2^3 - t_1^3\right)\right]^{1/\theta}$ $\theta = \alpha + \gamma(Gt/Gb)/tt$	-1.2879 (0.0854)	0.3418 (0.1886)	5.4039 (1.0715)	0.1081 (0.0637)
P. radiata (Eq. 9.10)	$N_2 = \left[N_1^{\alpha} + \frac{\beta}{1000000}\left(t_2^{\gamma} - t_1^{\gamma}\right)\right]^{1/\alpha}$	-1.6112 (0.0683)	0.0133 (0.00827)	2.1439 (0.0844)	- -

Notes: W=1 for Whirinaki Forest, 0 otherwise. Variables in parameter θ as defined in Section 9.4.2.

Only for *P. menziesii* was it possible to incorporate additional explanatory variables in the stocking model, which was based on an equation proposed by Clutter and Jones (1980). For this species the same thinning term as for the basal area model was introduced, along with altitude. As expected, thinning reduced mortality. The dummy variable for the Whirinaki forest was incorporated to model the increased mortality observed for that forest. Mortality was higher at lower altitudes, which may be caused by a greater incidence of *Phaeocryptopus* (promoted by the higher temperatures) and greater growth rates. The effect of thinning and altitude seemed to interact, incurring more mortality in unthinned stands at low altitudes than at higher altitudes (Figure 9.19). This would explain the same interaction observed in the basal area model (Figure 9.7).

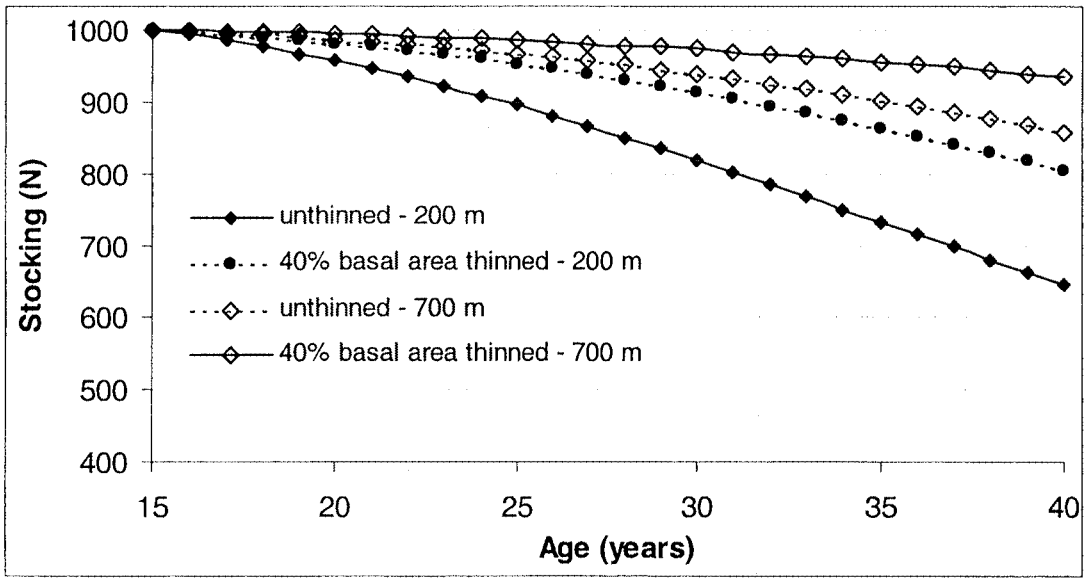


Figure 9.19 Projected stocking for unthinned stands and stands thinned to 40% of basal area at age 15 for two altitudes (*P. menziesii*).

9.6 MODELLING STAND VOLUME

In diameter distribution models, stand volumes can be explicitly obtained by summing up tree volumes across diameter classes weighted by frequency. However, if diameter statistics (i.e. maximum, standard deviation) are not available, the model can still be executed as a whole (or average) stand model, for which a stand volume equation would be required. Furthermore, even when volumes by diameter class can be predicted, some

researchers have preferred to have an explicit stand volume model and reconcile the two estimates (e.g. Woollons and Hayward 1985; Zhao 1999).

Typically, stand volume equations are static models dependent on predicted values of basal area and height (e.g. Woollons and Hayward 1985; Temu 1992), although stocking (Garcia, 1984) and age (Amateis *et al.* 1986) have also been used as independent variables.

A number of models representing well-established relationships between basal area, MTH and stand volume were tried (e.g. the combined variable model, $V=\alpha G^{\beta} MTH^{\gamma}$). However, the best fit was always achieved by exponential regressions of the following form:

$$V=\exp(a_0+a_1X_1+\ldots a_nX_n) \qquad (9.11)$$

where a_i are parameters and X_i independent variables including interactions and transformations.

Models were initially fitted and compared using unweighted regression, but final parameter estimates were obtained by weighted regression. The weighting factor was 1/(G.MTH) as the variance of residuals increased consistently with the product of basal area by MTH. Models fitted by weighted regression were more accurate (lower mean residual) than their counterparts fitted by unweighted regression (Table 9.11). This gain in accuracy was at the expense of a slight decrease in precision (higher standard deviation of residuals).

Table 9.11 Residual statistics for stand volume models.

	E. grandis		*P. menziesii*		*P. radiata*	
Residual statistics	unweighted	weighted	unweighted	weighted	unweighted	weighted
n	1023		835		1594	
Mean	-0.033	0.000	-0.043	0.000	-0.109	0.003
Standard deviation	2.58	2.59	9.20	9.21	10.78	10.80
Skewness	-0.12	-0.21	-0.25	-0.26	0.00	-0.01
Min (1%)	-7.2	-7.5	-25.5	-25.7	-34.7	-33.5
Max (99%)	7.4	7.3	25.7	25.5	30.2	31.7

Selected variables and parameter estimates for each species are presented in Table 9.12. For all species the logarithm of the product of basal area by MTH was the term that explained most of the variation in stand volume.

Table 9.12 Selected variables, parameter estimates and standard errors (in brackets) for stand volume models (Equation 9.11)

	Parameter estimates (standard errors)				
Species	a_0	a_1	a_2	a_3	a_4
E. grandis	**intercept**	**ln(G.MTH)**	**MTH/QMD**	**ln(t)**	-
	-0.8113	0.9963	-0.0732	-0.0205	-
	(0.00718)	(0.00209)	(0.00446)	(0.00336)	-
P. menziesii	**intercept**	**ln(G.MTH)**	**ln(N.MTH)**	**ln(SI)**	**(G.t)/1000**
	-1.0012	0.9394	0.0215	0.0563	0.00828
	(0.0268)	(0.00275)	(0.00125)	(0.00721)	(0.00138)
P. radiata	**intercept**	**ln(G.MTH)**	**1/t**	**ln(N.MTH)**	**ln(N)**
	-1.3667	0.9013	0.7414	0.1705	-0.1135
	(0.0308)	(0.0042)	(0.1158)	(0.0102)	(0.00896)

where QMD: quadratic mean diameter; SI: site index and other variables as previously defined.

Plots of residuals against predicted values are shown in Figures 9.20 to 9.22. Residuals show no trends, excepting the increasing variance for higher predicted values. Although this problem (called heteroscedasticity) was slightly lowered by using weighted regression, it could not be avoided completely. Its practical implications are unlikely to be of any importance.

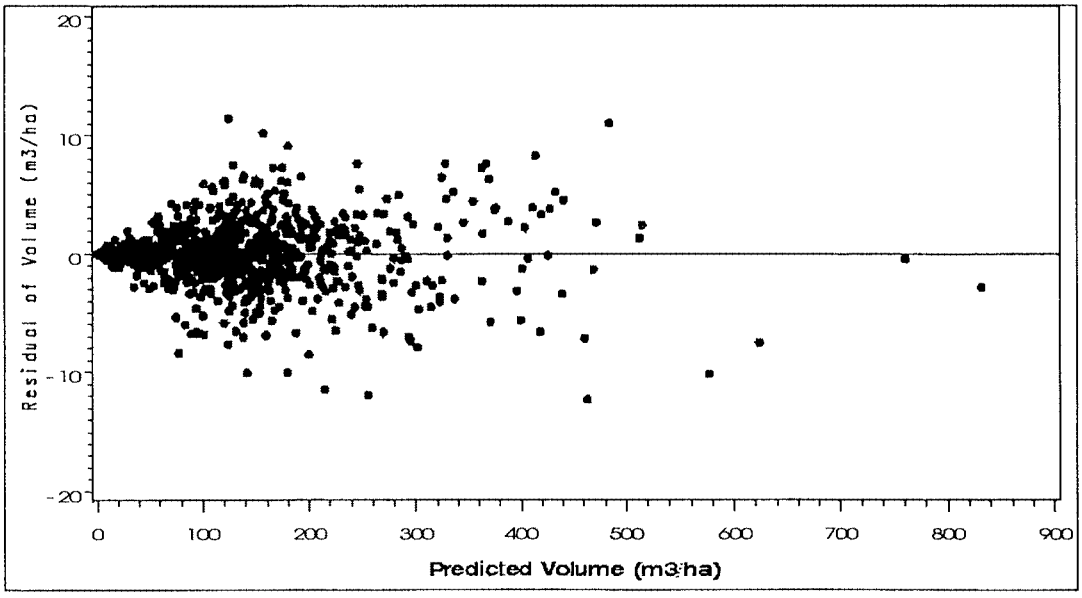


Figure 9.20 Residuals against predicted volumes in m^3/ha (*E. grandis*)

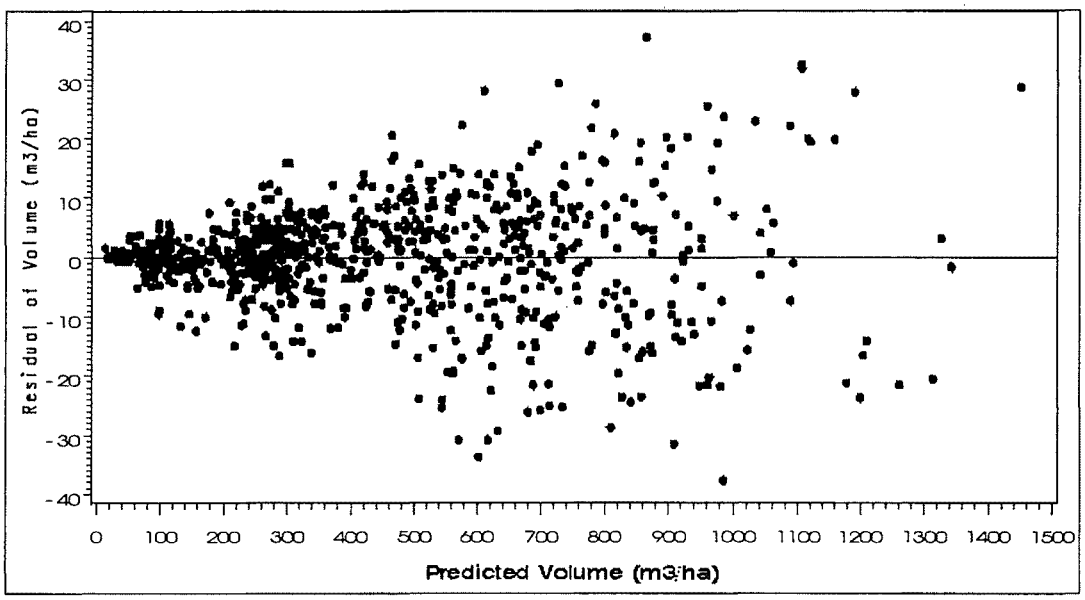


Figure 9.21 Residuals against predicted volumes in m^3/ha (*P. menziesii*)

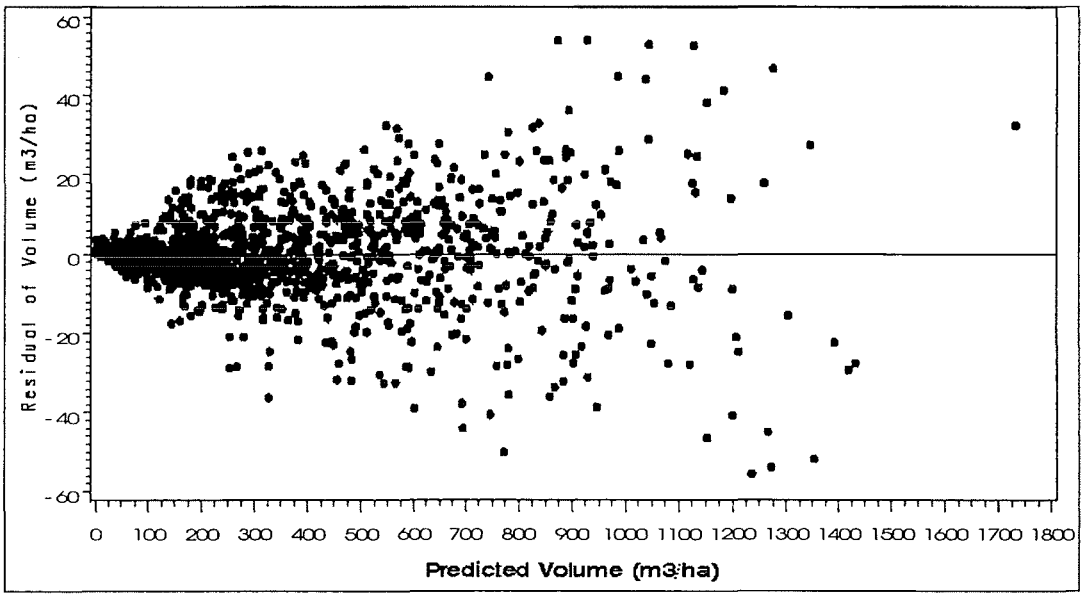


Figure 9.22 Residuals against predicted volumes in m^3/ha (*P. radiata*)

9.7 Derivation of Reverse Weibull Parameters

Having estimates of basal area, stocking, maximum diameter and standard deviation of diameters, the three parameters of the reverse Weibull distribution can be easily solved by the method of moments.

Theoretically, the maximum diameter constitutes the location parameter (a) of the reverse Weibull. However, as the *plot* maximum diameter may be lower than the *stand* maximum diameter (see Section 7.7.1 or García 1991), projected maximum diameters should be increased somehow. Xu (1990) and Xu *et al.* (1992) suggested using extreme distribution theory for that purpose, although their approach may have shortcomings (see Section 7.7.1). A simple alternative strategy consisting of increasing maximum diameters estimated with equations 9.3 to 9.5 by a fixed percentage was explored. Stand tables were generated using unadjusted predicted maximum diameters as well as predicted maximum diameters increased by 5, 10, 15 and 20%. Error indices (Reynolds *et al.* 1988) were calculated with various weighting factors (i.e. tree volume, tree basal area and no weighting factors). The best overall results were achieved by increasing the projected maximum diameters by 10, 15 and 20% for *P. menziesii*, *P. radiata* and *E. grandis*, respectively. It must be noted that predictions of maximum diameter (D_{max}) are independent of calculations of quadratic mean diameters (QMD) predicted from basal area and stocking. When using the model for very low stand densities, it may be possible (although rare) for QMD to exceed D_{max}. In order to prevent this situation, the value of D_{max} can be constrained to be larger than QMD by a certain minimum percentage. Minimum percentages observed in the datasets were 12.6, 7.6 and 2.0% for *E. grandis*, *P. menziesii* and *P. radiata*, respectively.

The reverse Weibull parameters were obtained as follows:

$a=\hat{D}_{max}*1.15$ (the factor 1.15 applies for *P. radiata*)

$Variance=\hat{D}_{std}^{2}$

$QMD=\sqrt{\frac{\hat{G}40000}{\hat{N}\pi}}$ (QMD: quadratic mean diameter)

$MD=\sqrt{QMD^{2}-Variance}$ (MD: arithmetic mean diameter)

$z=D_{std}/(a\text{-}MD)$

$kz0=-0.22004032*z^0$

$kz1=-0.001433169*z^1$

$kz2=0.150611381*z^2$

$kz3=-0.078575996*z^3$

$kz4=0.004305716*z^4$

$kz5=0.008804944*z^5$

$$c=\frac{1}{z\left[1+(1-z)^2(kz0+kz1+kz2+kz3+kz4+kz5)\right]}$$

$$b=\sqrt{\frac{\text{Variance}}{\Gamma\left(1+\frac{2}{c}\right)-\Gamma^2\left(1+\frac{1}{c}\right)}} \quad \text{or} \quad b=\frac{a-MD}{\Gamma\left(1+\frac{1}{c}\right)}$$

The above steps were based on García's (1981) simplified method of moments, although they were adapted to the reverse Weibull distribution (note that García's notation for the Weibull distribution is different from the more usual notation, as in Section 7.7.1).

9.8 VERIFICATION AND VALIDATION

Some verification of model logic for individual components was already performed in the previous section (e.g. plots of projected basal area and MTH at varying altitude, zone and thinning intensity). In this section, the performance of the models is further evaluated in three ways:

(i) comparing projections of some components (MTH, G) against projections obtained with previous models;

(ii) verifying the performance of each component through analysis of residual statistics computed from validation subsets; and

(iii) predicting diameter distributions with the whole modelling systems (i.e. integrating all components) and comparing them to actual stand tables.

9.8.1 Comparison with previous models for some individual components

9.8.1.1 *E. grandis*

For *E. grandis* in Uruguay there are no growth and yields models available. However, simple site index models have been constructed in the country (Sorrentino 1992) and in the neighbouring province of Entre Ríos, Argentina (Glade 1999). Although, the mentioned models were derived through algebraic manipulation of yield-type data of MTH, they can be converted into projection equations.

Mean top heights of all plots were projected from the first measurement to the last measurement for existing site index models and for the model developed here. Plots of residuals against predicted values are displayed in Figure 9.23.

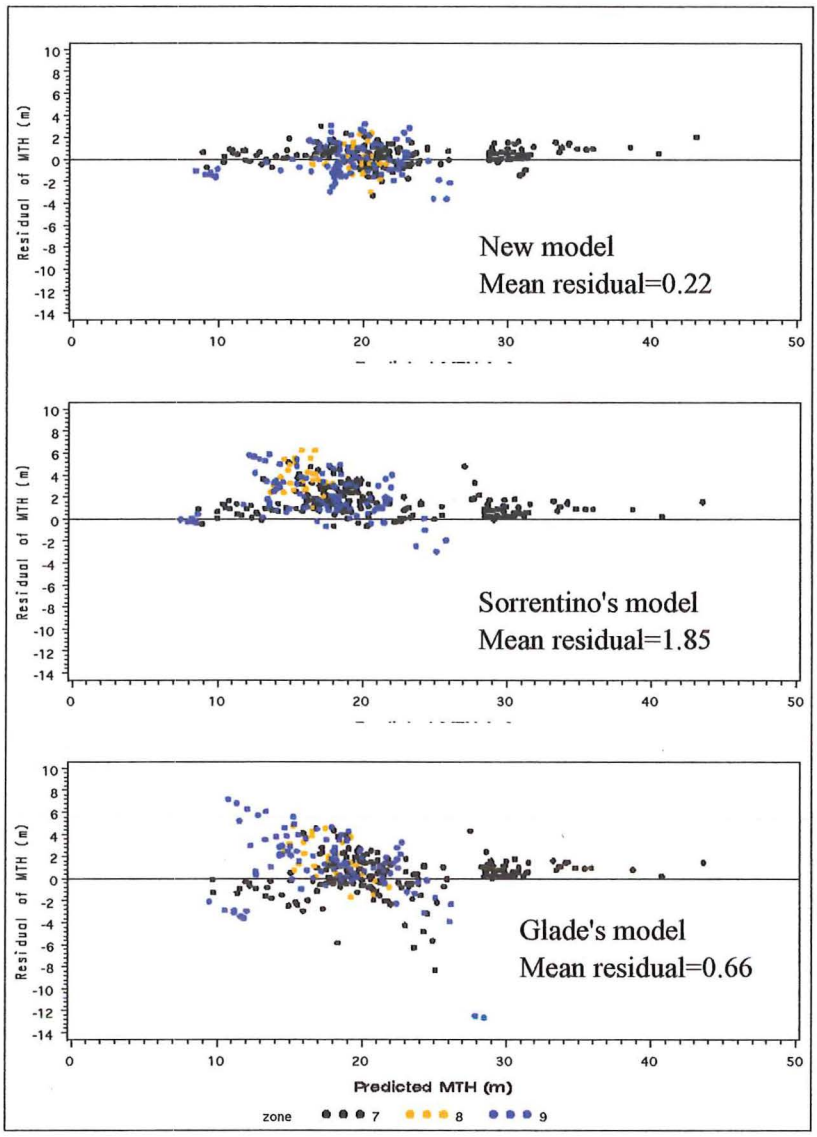


Figure 9.23 MTH residuals against predicted values (m) for the new model (top), Sorrentino's model (middle) and Glade's model (bottom).

Although the comparison is not completely fair, because the data used in the comparison were included in the fitting dataset, residual trends show unambiguously the better performance of the new model. This could be attributed to the larger amount of data available for this study plus the use of the difference equation method (as opposed to the guide curve method) for developing the site index models.

9.8.1.2 *Pseudotsuga menziesii* (Douglas-fir)

For this species basal area and MTH models developed here were compared to those developed by Xu (1990) for the same region. Only measurements taken after 1966 from

the Kaingaroa forest were used (Xu's MTH model has different sets of parameters for different forests). All intervals not used for model fitting were used in the comparison, thus providing some degree of independence from the fitting dataset. Plots of MTH residuals against predicted values and against interval length (years) are shown in Figures 9.24 and 9.25 respectively.

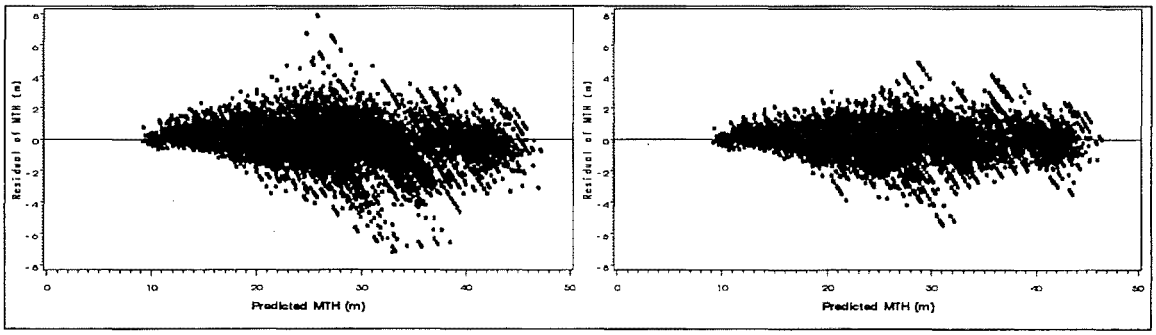


Figure 9.24 MTH residuals against predicted values from Xu's model (left) and the new model (right).

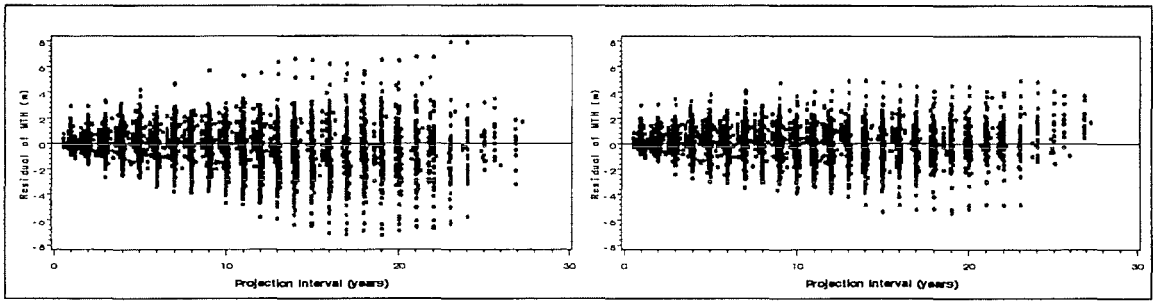


Figure 9.25 MTH residuals against interval length (years) from Xu's model (left) and the new model (right).

The distribution of the residuals from the new model is much tighter than the residual distribution from Xu's model. The mean residual was -0.042 and -0.196 for the new model and Xu's model, respectively. All other residual statistics were consistently better for the new model. By analysing the trend of residuals with interval length (Figure 9.25), it can be noted that the superiority of the new model was more evident over long projections. This, in turn, was probably due to the use of mixed intervals (including long intervals) for fitting the new model, whereas only consecutive intervals were used by Xu. Another possible cause of the better fit of the new model is that the selected model form was a polymorphic Chapman-Richards equation, whilst Xu's model was an anamorphic

version of the same model. Polymorphic equations are often more flexible than anamorphic equations for representing height and basal area development.

Xu's basal area model lacks the property of path-invariance requiring annual iterations for its implementation. Therefore, only consecutive intervals (from those not used for fitting the new model) were selected for the comparison. Basal area residuals are plotted against predicted values in Figure 9.26.

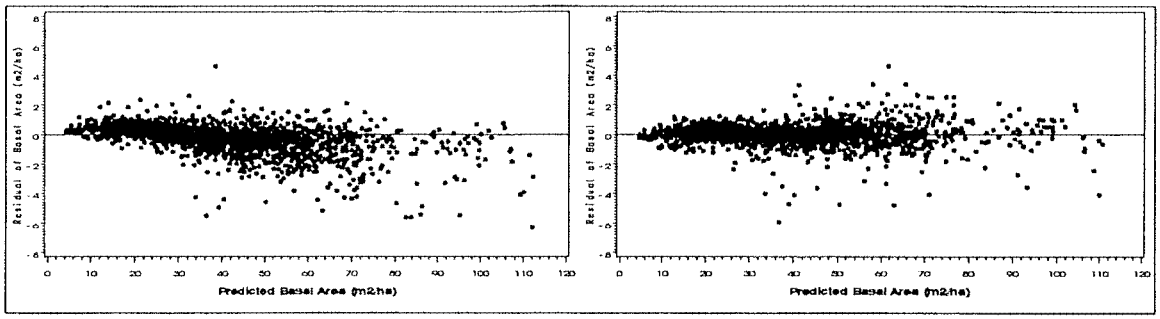


Figure 9.26 Basal area residuals against predicted values from Xu's model (left) and the new model (right).

From Figure 9.26 it appeared that larger predicted values were consistently overpredicted with Xu's model, whereas the new model showed no patterns. The better performance of the new model for basal area projections may be explained by the inclusion of altitude and a different thinning index coupled with thinning age in the new model. Another difference between the two basal area models was the selected equation form. While a Schumacher-based polymorphic equation, which was proven the best equation form for the three species, was used in this study, a Hossfeld-type equation was preferred by Xu (1990). Additionally, the new model was fitted to a larger dataset containing 10 years of new measurements, which may provide more information on the asymptotic region of basal area growth.

9.8.1.3 *P. radiata*

The current stand level model for this species in the Kaingaroa region is PPM88, which was fitted using García's (1984, 1994) state-space approach. This means that all variables were fitted simultaneously through a system of stochastic differential equations. Instead

of age, MTH represents the time scale in PPM88 and, therefore, the MTH model is a key component of the system.

The MTH model in PPM88 is the same as in KGM3, a polymorphic Chapman-Richards equation. The MTH model developed in this study is also a polymorphic Chapman-Richards equation but derived from another parameter, as illustrated in Section 7.4.3.2. Mean top height projections from PPM88 and from the new model were made for the validation dataset (plots not used for model fitting). Residual plots are displayed in Figure 9.27.

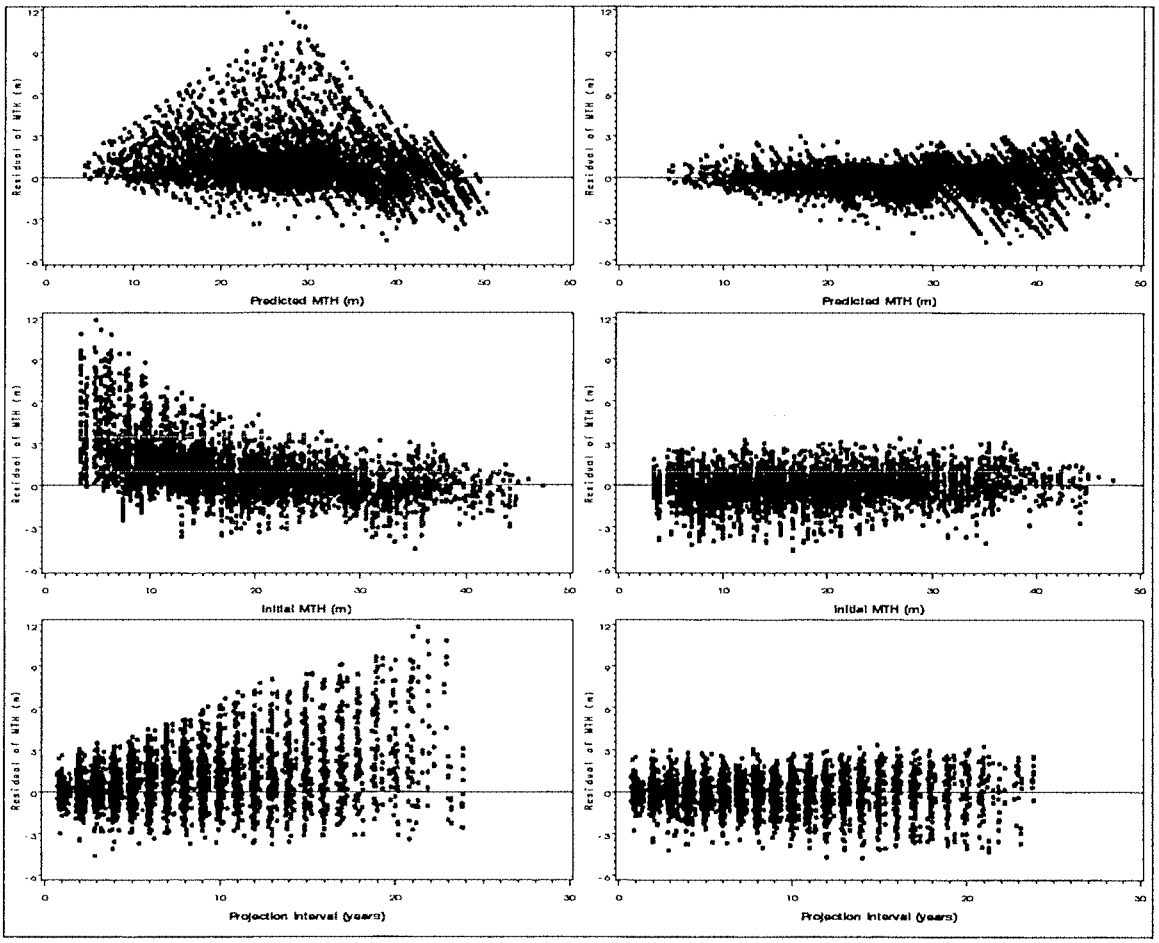


Figure 9.27 MTH residuals against predicted values (top), initial MTH (middle) and interval length (bottom), from PPM88 (left) and the new model (right).

MTH residuals from PPM88 suggested a highly biased model at a first glance. Analysing in detail the three plots in the left-hand side of Figure 9.27, it can be noted that most extreme residuals appeared when plots with low MTH were projected over long intervals. The documentation of the PPM88 software (Dunningham, A.G.; García, O. and

M.E. Lawrence, February 1989) indicated that "the model may not be reliable for stands under 8 m top height". This explains why PPM88's MTH model performed so poorly for initial top heights lower than 8 m, although the zone of positive residuals (underprediction) continues up to approximately 15 m of initial MTH. The fact that PPM88's MTH model performed consistently poorer over increasing projection lengths may be a consequence of the model being fitted to consecutive intervals only. The new model, on the other hand, was fitted to mixed-length intervals and exhibited little bias over projections up to 24 years, with 90% of the residuals between -2.00 and +1.65 m.

Other possible reasons for the better performance of the new MTH model include (i) using a different form of the Chapman-Richards model (which was found best for the three datasets); (ii) incorporating the altitude effect and (iii) being fitted to a larger dataset.

PPM88, which is probably the most used stand model for *P. radiata* in the Kaingaroa region, works under the assumption that basal area and stocking vary consistently with MTH, rather than with time. The relatively poor performance of its MTH model for young stands over long projections may jeopardise the performance of the whole system for these situations.

Mason *et al.* (1997) found a good linkage between PPM88 and the initial growth model (IGM) for the same region that incorporates silvicultural treatments applied in the establishment phase (Mason 1992; Mason and Whyte 1997). However, the poor performance of PPM88's MTH model for young stands suggests that the linkage between IGM and growth and yield models for older crops could be (potentially) substantially improved.

Basal area projections with PPM88 would require information on closure (García 1990) and green crown level. These data were not available for this study. Making assumptions about these variables would have provided a non-optimal use of PPM88, and an unfair comparison against the model developed in this study. Therefore, a full comparison of both models was not performed but it may be warranted in further studies.

9.8.2 Validation of main components

For the New Zealand-grown species, data subsets consisting of independent intervals (*P. menziesii*) or plots (*P. radiata*) were set aside for validation purposes (as explained in Section 8.4.1). The developed models were applied to validation datasets, and residuals were computed and summarised in tabular form.

The validation of the *E. grandis* model components was basically limited to the tables and graphs of residuals shown in Sections 9.4 to 9.6. However, stand volume predictions were further evaluated in this section by using the stand volume model with predicted values of independent variables.

Residual statistics for MTH, basal area/ha, maximum diameter and standard deviation of diameters, from validation subsets are displayed in Table 9.13 (stocking residuals for validation subsets were already shown in Table 9.9).

Table 9.13 Residual statistics of MTH, basal area, D_{max} and D_{std} (validation subsets).

	P. menziesii (n=12448)				*P. radiata* (n=5786)			
Residual statistic	MTH (m)	G (m^2/ha)	D_{max} (cm)	D_{std} (cm)	MTH (m)	G (m^2/ha)	D_{max} (cm)	D_{std} (cm)
Mean	-0.04	0.01	0.20	-0.04	-0.13	0.50	0.51	0.05
Standard deviation	0.94	2.35	1.97	0.69	1.10	3.57	3.46	0.97
Skewness	-0.11	0.28	0.65	-0.74	-0.33	0.21	0.06	-0.48
Minimum (1%)	-2.61	-6.21	-4.53	-2.56	-3.37	-9.74	-8.30	-2.89
Maximum (99%)	2.39	7.13	5.83	1.92	2.41	11.01	10.03	2.61

The overall accuracy of all model components (measured through mean residuals) was satisfactory. The largest mean residuals corresponded to basal area and maximum diameter of *P. radiata* and were still reasonable in relative terms. The standard deviation of residuals, which is a measure of the overall precision of the models, was greater for basal area. This was expected, as basal area growth is more affected by stocking and silvicultural treatments (thinning, pruning), and thus is more difficult to model than MTH for instance. Additionally, the larger standard deviations for basal area and maximum diameter are also due to their large magnitude in absolute terms. The distribution of residuals, ascertained by the skewness statistic and the balance between maxima and minima, was largely satisfactory.

For the stand volume equation, residuals were calculated from (i) actual values of the independent variables (i.e. MTH, G, N) and (ii) predicted values of the independent variables. As expected, residual statistics worsened when predicted values were used (Table 9.14) as errors of all components were compounded.

Table 9.14 Residual statistics of stand volume (m^3/ha) computed to validation subsets.

	E. grandis		*P. menziesii*		*P. radiata*	
Residual statistic	estimation*	prediction	estimation	prediction	estimation	prediction
n	1023	709	1524	12448	1237	4874
Mean	0.00	-0.09	0.72	-0.18	1.25	1.29
Standard deviation	2.59	15.21	9.72	30.92	8.91	44.76
Skewness	-0.21	-0.24	-0.59	-0.05	1.53	0.90
Minimum (1%)	-7.5	-37.4	-27.7	-81.7	-16.7	-116.5
Maximum (99%)	7.3	35.5	27.7	89.6	35.7	170.4

* Estimation: volumes calculated from *actual* stand variables (basal area, MTH and stocking). Prediction: volumes calculated from *predicted* stand variables.

For *E. grandis* and *P. menziesii*, volume residual statistics showed little bias. For *P. radiata*, residuals were unbalanced towards positive values (underestimation) either when actual (estimation) or predicted (prediction) values for required stand variables were used. Stocking and basal area were both slightly underestimated for this species (Tables 9.9 and 9.13) and these errors were compounded in the stand volume model. Extreme positive residuals corresponded to the highest predicted values, indicating that most of the overall underestimation occurred for extremely high volumes, which are unlikely to be achieved in practice.

9.8.3 Stand-table projection performance

There are several statistical tests available to ascertain the degree of concordance between estimated and actual diameter distributions. These include the Pearson chi-square test, the Kolmogorov-Smirnov (KS) test, the Cramér-von Mises (CM) test and the Anderson-Darling statistic (Reynolds *et al.* 1988). These tests aid to determine whether to accept or reject the null hypothesis that actual and predicted distributions are equal. Whilst this type of test may be reasonable from a statistical viewpoint, it provides limited information regarding the validity of the models for practical applications. Furthermore,

after a thorough evaluation of various goodness-of-fit tests Reynolds *et al.* (1988) concluded that they may be inappropriate for diameter distribution models. Those authors proposed an error index that overcomes most of the shortcomings of traditional statistical goodness-of-fit tests (see Chapter 12). Although this method is valid for comparing models, it does not provide an interpretable stand-alone measure of the quality of a single model.

Some researchers used frequency plots of actual and predicted distributions for individual plots to illustrate the degree of accuracy of the model (e.g. Mason and Whyte 1997; Frazier 1981). This approach was also adopted here. Eight intervals from eight plots were randomly selected for each species, and actual and predicted diameter distributions were plotted (Figures 9.28 to 9.30). The validation plots for only *P. radiata* were different to the plots used for model fitting. For *P. menziesii*, selected intervals were not used for fitting any of the model components (although other intervals from the same plots were).

For *E. grandis*, projected diameter distributions were very close to actual ones for projections up to 6 years long. For *P. menziesii* and *P. radiata*, as projection lengths of the selected intervals increased, the degree of agreement between actual and predicted distributions was slightly inferior. The main reason for that was probably the difference between estimated and actual stockings (shown in each plot), although basal area, maximum diameter and standard deviation errors also contributed. The accuracy of maximum diameter predictions (which determine the location parameter of the Weibull distribution) was a crucial factor for achieving accurate diameter distributions. The two plots on the bottom of Figure 9.30 illustrate this (for plot 9501 maximum diameter was overestimated whereas for plot 9350 maximum diameter was underestimated).

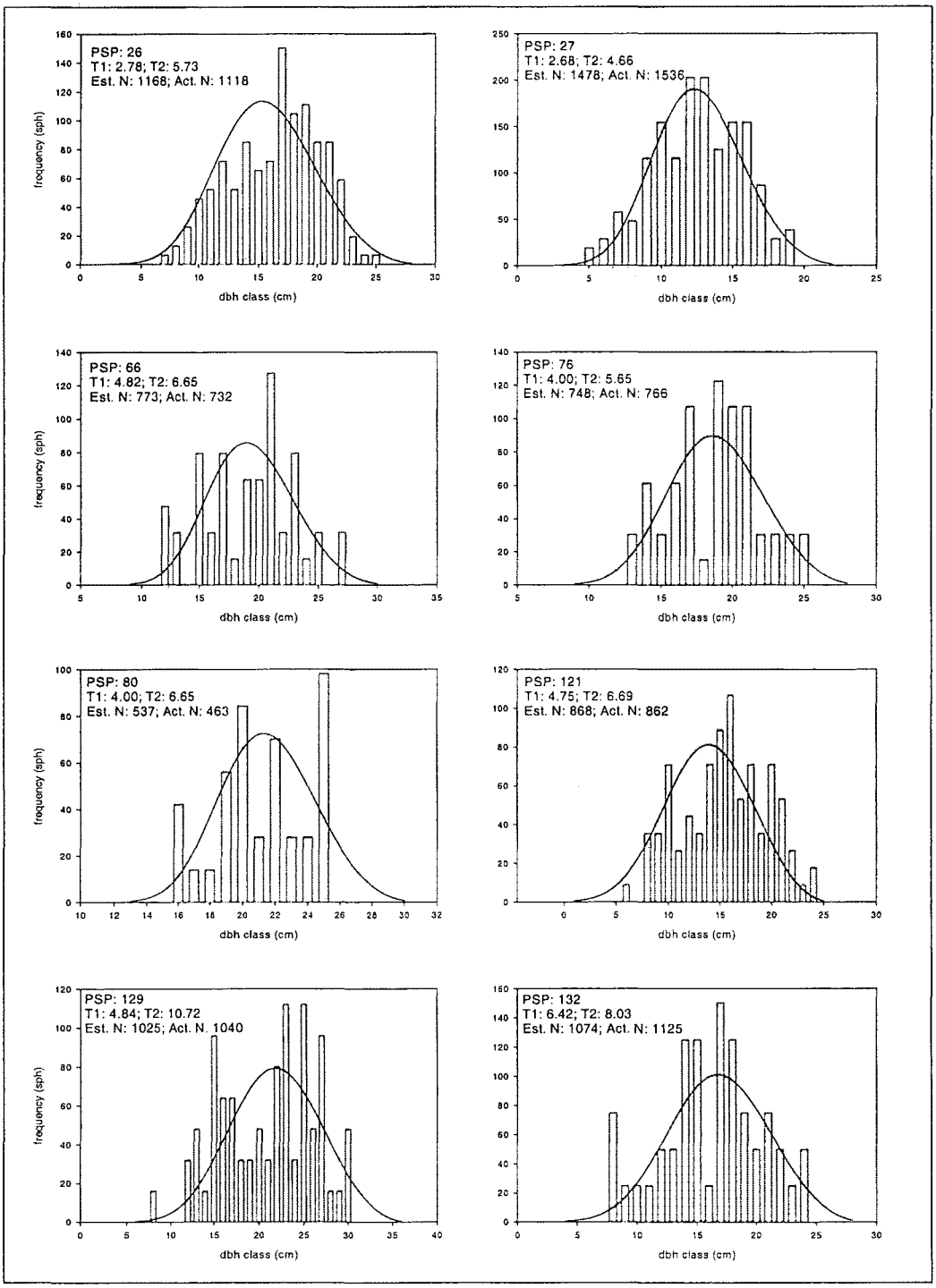


Figure 9.28 Projected *versus* actual diameter distributions (*E. grandis*).

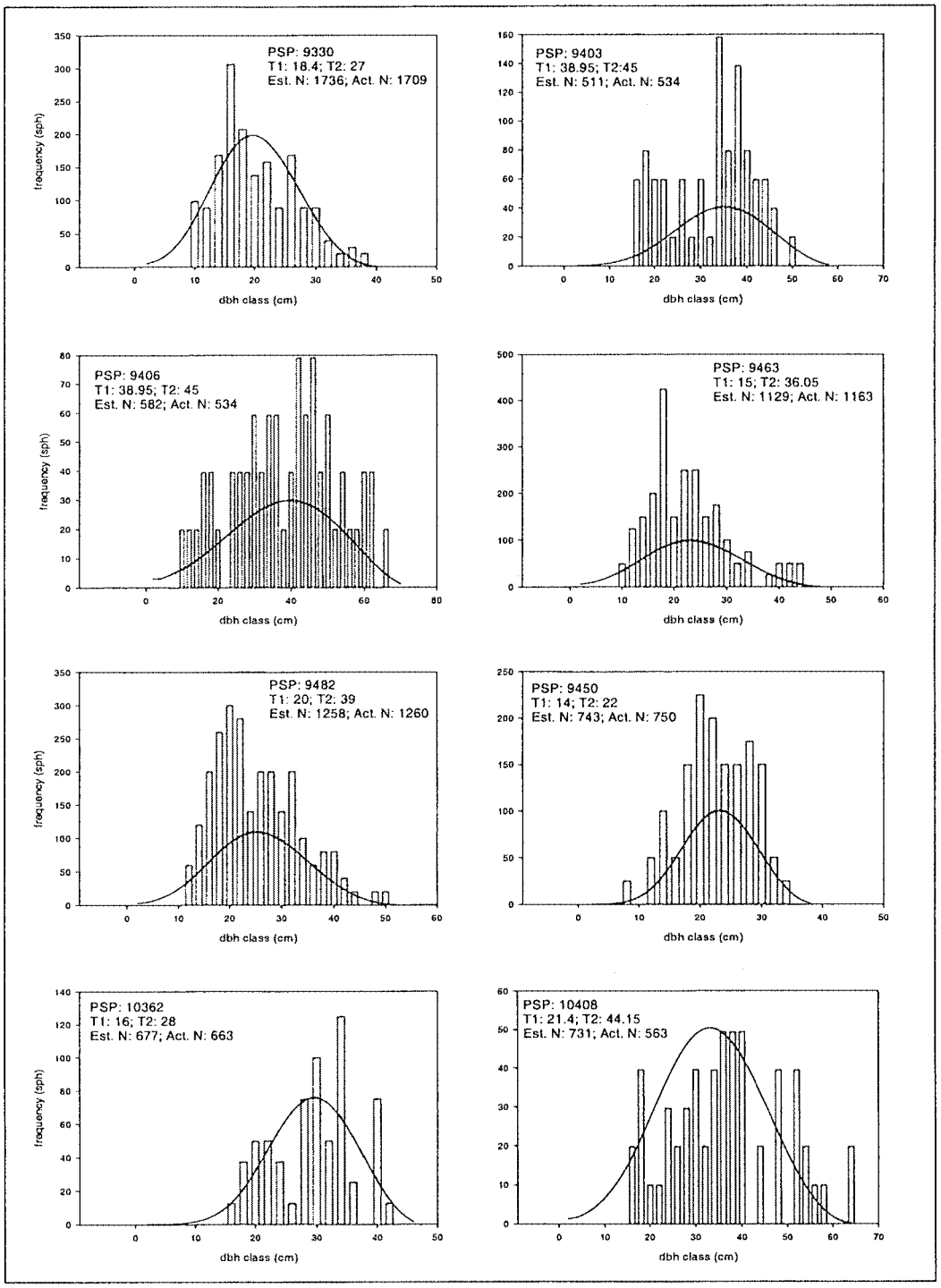


Figure 9.29 Projected *versus* actual diameter distributions (*P. menziesii*).

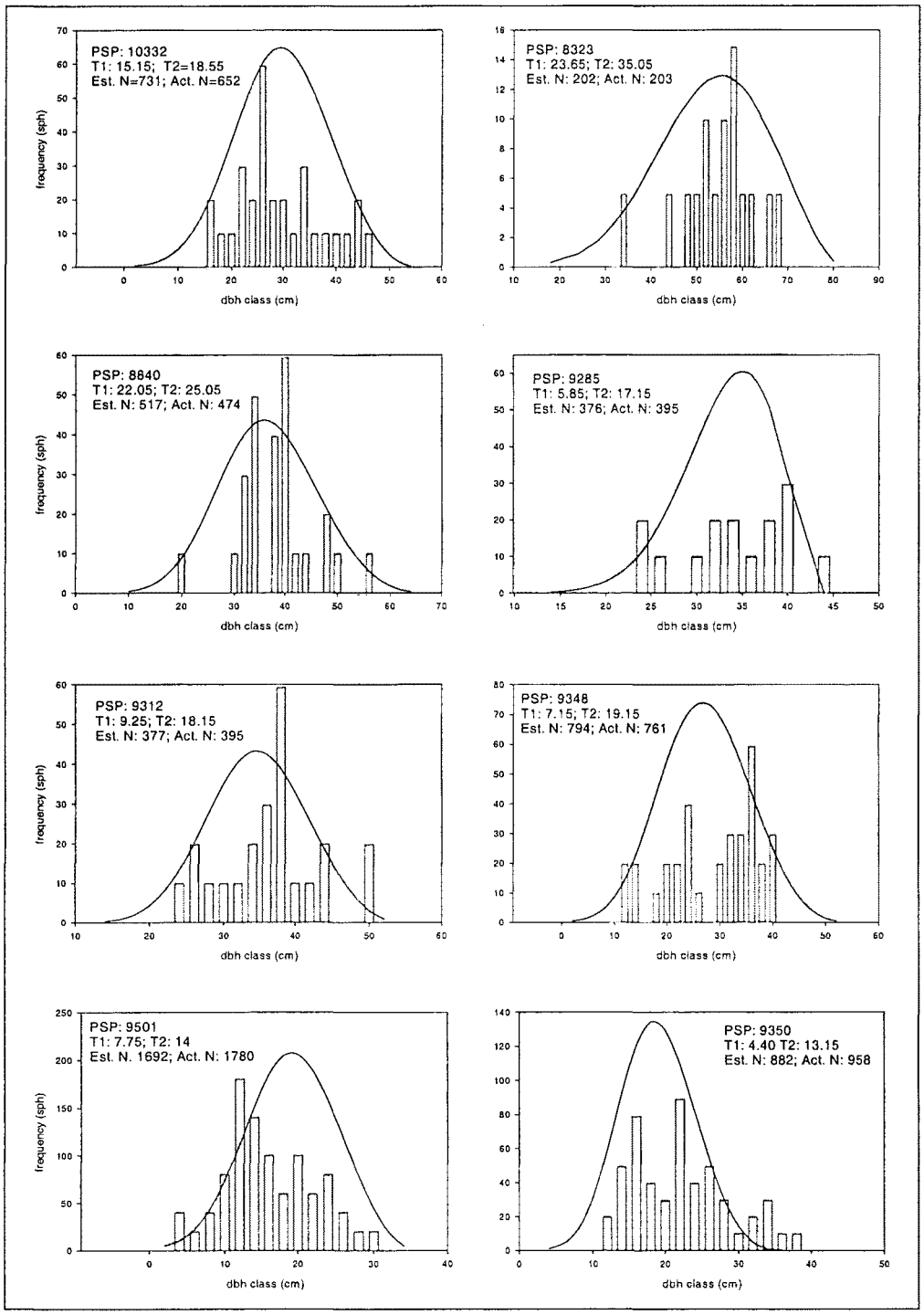


Figure 9.30 Projected *versus* actual diameter distributions (*P. radiata*).

The reasonable level of agreement between predicted and actual distributions for most plots in Figures 9.28 to 9.30 confirmed the good performance of the reverse Weibull parameter-recovery approach advocated by Whyte and Woollons (1992).

9.9 BASAL AREA AND STANDARD DEVIATION AFTER THINNING

For implementation of the developed models, two further components were needed. These are models that simulate the instant change in (i) basal area/ha and (ii) standard deviation of diameters as a result of a thinning operation. Maximum diameter is hypothesised to remain the same, provided that thinnings are from below.

These models were developed from all thinning events available in each dataset and represented thinnings from below. For *P. radiata* only there were some few thinnings from above but they were excluded from the analysis.

The selected models with corresponding parameter estimates are shown in Tables 9.15 and 9.16. All models performed satisfactorily showing no trends with any of the independent variables. Statistics of residuals from these models are shown in Table 9.17.

Table 9.15 Models for predicted basal area/ha after thinning

Species	Model	Parameter estimates (standard errors) α	β	γ	δ
E. grandis (Eq. 9.12)	$G_a = \alpha G_b^{\beta}\left[1-\left(1-\frac{N_a}{N_b}\right)^{\gamma}\right]^{\delta}$	1.1592 (0.0538)	0.9148 (0.0157)	2.2221 (0.1517)	1.2664 (0.0702)
P. menziesii (Eq. 9.13)	$G_a = \alpha G_b^{\beta}\left[1-\left(1-\frac{N_a}{N_b}\right)\right]^{\delta}$	1.1015 (0.0674)	0.9759 (0.0148)	- -	0.6786 (0.0127)
P. radiata (Eq. 9.14)	$G_a = \alpha G_b^{\beta}\left[1-\left(1-\frac{N_a}{N_b}\right)^{\gamma}\right]^{\delta}$	0.9382 (0.0387)	1.0237 (0.00996)	0.9695 (0.0578)	0.7459 (0.0190)

Notes: G_b and G_a are stand basal areas (m^2/ha) before and after thinning and N_b and N_a are stockings (stems/ha) before and after thinning.

Table 9.16 Models for standard deviation after thinning

Species	Model	Parameter estimates (standard errors) α	β	γ
E. grandis (Eq. 9.15)	$STD_a=STD_b\exp\left[\left(\alpha+\beta STD_b^2\right)\left(\frac{N_b-N_a}{N_b}\right)^{\gamma}\right]$	-0.3590 (0.1186)	-0.0763 (0.0179)	1.1531 (0.1572)
P. menziesii (Eq. 9.16)	$STD_a=\alpha STD_b^{\beta}\left[1-\left(1-\frac{N_a}{N_b}\right)\right]^{\gamma}$	0.9125 (0.0627)	1.0044 (0.0300)	0.1784 (0.0224)
P. radiata (Eq. 9.17)	$STD_a=\alpha STD_b^{\beta}\left[1-\left(1-\frac{N_a}{N_b}\right)\right]^{\gamma}$	1.0250 (0.0424)	0.9314 (0.0180)	0.1710 (0.0174)

Notes: STD_b and STD_a are standard deviations of diameters (cm) before and after thinning

Table 9.17 Statistics for residuals from models of basal area and standard deviation of diameters after thinning.

	Basal area after thinning			Standard deviation after thinning		
Residual statistics	*E. grandis*	*P. menziesii*	*P. radiata*	*E. grandis*	*P. menziesii*	*P. radiata*
Mean	0.00	0.07	0.05	0.00	-0.01	0.01
Standard deviation	0.26	2.19	1.30	0.24	0.72	0.66
Skewness	-0.16	0.49	0.32	0.39	-0.01	-0.11
Minimum (1%)	-0.87	-4.94	-3.26	-0.45	-2.06	-1.95
Maximum (99%)	0.69	6.27	4.00	0.60	1.61	1.58
Number of observ.	62	161	549	58	142	385

9.10 Discussion and Conclusions

Mean top height and basal area/ha were best modelled with the same equation form for all three species. The selected model for dominant height was a polymorphic Chapman-Richards equation. Although other equations derived from the Chapman-Richards model have been found appropriate for height modelling by a number of researchers (e.g. García 1981, 1988a; Pienaar 1989, Candy 1997, Martin *et al.* 1999), the polymorphic variant selected here has seldom been used. An exception is the study by Amaro *et al.* (1998) wherein the same polymorphic variant was found best for *Eucalyptus* plantations in Portugal. The flexibility of this model is such that the projected curves may or may not have a lower inflection point. When there is no lower inflection point (as observed for many plots in the *E. grandis* dataset) the initial growth is rapid and the instantaneous growth rate decreases monotonically from age zero onward.

A negative altitudinal effect on dominant height growth was readily apparent for *P. radiata* and the introduction of altitude in the model improved the fit markedly. Although the effect of altitude on height growth of *P. menziesii* was lower, it was incorporated in the model as it provided a slightly better fit and because altitude information is easily available.

The selected basal area/ha model was the well-known 2-parameter Schumacher polymorphic model, which has been consistently chosen in many studies for basal area modelling (e.g. Clutter and Jones 1980; Woollons and Hayward 1985; Kuru 1989; Temu 1992; Candy 1997; Mason and Milne 1999; Zhao 1999). Several earlier studies have introduced site index in the basal area equation to account for differences in site productivity (e.g. Schumacher 1939; Clutter 1963; Bailey and Ware 1983; Murphy and Farrar 1988). For the New Zealand datasets, altitude was a better predictor variable than site index and completely subsumed any impact that site index might have on basal area growth. This is consistent with results from other modelling studies for the same species growing in New Zealand (Temu, 1992; Whyte *et al* 1992; Lee 1998; Zhao 1999). For *E. grandis*, stand basal area growth was better adjusted through dummy variables for locality than through site index.

In addition to altitude, the effect of thinning was incorporated in stand basal area models. Thinning terms modified the growth of thinned stands relative to the growth of unthinned stands with identical basal area/ha at the same age. For *P. radiata*, basal area growth of the thinned stands was found lower than basal area growth of unthinned stands. This is consistent with results from other modelling studies that found that the loss of crown surface by thinning and pruning operations reduced basal area growth (West *et al.* 1982; West *et al.* 1987; García 1990). The effect of pruning on basal area growth could not be properly modelled with the data available for this study. Pruning operations were often performed at the same time as thinning operations and, therefore, the reductions in basal area growth after thinning found for *P. radiata* may be exaggerated by the loss of crown surface through pruning. Reductions in basal area yields at age 30 caused by thinning 50% of the basal area at age 5 were between 15.2 and 17.9% depending on the altitude level (see Figure 9.8).

For *P. menziesii*, the thinning effect on basal area growth was positive, indicating that if a stand were thinned to the same basal area/ha as another (unthinned) stand of identical age, it would grow at a faster rate than the unthinned stand. This effect may have been caused by the reduced mortality registered in thinned stands, particularly at lower altitudes. As all thinnings were from below, subdominant trees that are more prone to *Phaeocryptopus* were removed, thus reducing further mortality and increasing the vigour of the stands. Basal area yields at age 40 in thinned stands (40% of basal removed at age 15) were 4.9 and 11.0% higher than in unthinned stands, for altitudes of 700 and 200 m a.s.l respectively (see Figure 9.7). For the same conditions, mortality was higher in unthinned plots, especially at the lowest altitude (see Figure 9.19).

Selected forms of the stocking projection equations were variants of models proposed by Clutter and Jones (1980) and Clutter *et al.* (1983), for the New Zealand and Uruguayan datasets, respectively. Altitude and thinning indices were successfully introduced in the model for *P. menziesii* as explained in Section 9.5. Among the three approaches tried for fitting stocking models, using the longest interval of each plot provided the best overall results. This approach was found to be a simple and effective solution to some of the problems associated with mortality modelling (e.g. Woollons 1998).

Stand volume was best modelled through exponential regressions of the form $V=\exp(a_0+a_1X_1+\ldots a_nX_n)$, the variable explaining most of the variation being ln(MTH.G) for all three datasets. This model form has not been commonly adopted in growth modelling studies (an exception being Amateis *et al.* 1986), but allows incorporating important explanatory variables (e.g. stocking, age) that are usually neglected in traditional stand volume equations. Weighted non-linear regressions provided better models than unweighted non-linear regressions, with mean residuals virtually zero and an excellent balance between positive and negative residuals.

Projection equations for maximum diameter and standard deviation were built, thus completing the necessary elements for deriving the three parameters of the reverse Weibull distribution, through a method-of-moments parameter-recovery approach. Diameter distributions were projected using the entire system of equations and matched very well with actual stand tables, as illustrated in Figures 9.28 to 9.30.

Increasing predicted maximum diameters by a fixed percentage for computing the location parameter provided better results than using unadjusted estimates. Although this empirical solution to the problem of defining the population extreme performed satisfactorily (see Figures 9.28 to 9.30), further theoretical studies following the work by Kuru *et al.* (1992) and Xu *et al.* (1992) may be needed to clarify this topic. A potentially valid approach would be to develop a yield model of *stand* (as opposed to plot) maximum diameter on age and other explanatory variables (e.g. altitude or site index, and some measure of stand density) and use it coupled with a maximum diameter projection equation. Predictions obtained from the two models could be reconciled somehow to ensure that the location parameter properly represents the maximum stand diameter attainable at any given time.

Basal area/ha and MTH projections obtained from the models developed in this study performed better than projections from existing models. For *E. grandis* in Uruguay there were no complete growth and yield models available and, therefore, the diameter distribution model developed here can be used meanwhile. However, given that the model was fitted to a reduced dataset, it should be revised when more PSP data became available.

CHAPTER 10

INDIVIDUAL TREE MODELS

10.1 INTRODUCTION

Growth and yield estimations from average stand level models or even from diameter distribution models may not be sufficiently detailed to allow optimal management decisions to be made. Size information at the tree level can be represented by individual tree models, which have been traditionally categorised into distance-dependent and distance-independent (Munro 1974). Whether or not tree spatial information can improve tree-growth predictions for plantations has been extensively debated (Daniels and Burkhart 1975; Clutter *et al.* 1983; Bruce and Wensel 1987; Biging and Dobbertin 1995). The unavailability of inter-tree distances in the datasets available for this study, precluded the utilisation of the distance-dependent approach.

Complete individual tree models require three basic components, namely (i) a model to predict tree diameters (or tree basal areas), (ii) a model to predict the probability of tree survival, and (iii) a model to predict individual tree heights. The development of the former two components is reported in this chapter, whereas the latter component is discussed in Chapter 11.

Predictions of tree diameters can be achieved by developing diameter increment models based on, for instance, the tree's initial size, its competitive status or hierarchy within the

stand, and measures of stand density and site productivity. These factors have all been used in several diameter/tree basal area increment models (e.g. Alder 1979; Manley 1981; Shula 1997b; Murphy and Graney 1998; Lynch *et al.* 1999; Huebschmann *et al.* 2000). A potentially useful alternative would be to use traditional difference equations to explicitly model diameters at any given age based on initial diameter, initial and final ages and possibly other stand or site variables (e.g. site index, altitude). Masripatin (1998) found this approach better than the traditional diameter-increment approach for mixed tropical rain forests in Indonesia. This was because age of trees was not known and so the gap in time between successive measurements was used to represent (t_2-t_1).

Another approach for predicting tree diameters is to apportion predicted stand basal areas (obtained by stand level models) to individual trees by some dis-aggregating method. A dis-aggregating method that has demonstrated excellent potential is based on the relative size of the tree measured through the ratio of the tree's basal area to the average tree basal area (Clutter and Allison 1974; Pienaar and Harrison 1988). This approach ensures consistency between stand- and tree-level estimates, whereas such consistency is not intrinsically assured by diameter or diameter increment models (although it can be forced).

The probability of tree survival has been traditionally modelled by the logistic model. This approach has been successfully applied in many studies (e.g. Hamilton 1974; Monsreud 1976; Avila and Burkhart 1992; Shula 1997a; Murphy and Graney 1998; Huebschmann *et al.* 2000), being flexible and easy to interpret. This method was also used in the study reported here.

10.2 METHODS AND PROCEDURES

As explained in Section 8.4.2, two sets of tree-level intervals were constructed for each species, namely (i) non-overlapping intervals and (ii) mixed-length intervals. From each set, a single interval per tree was randomly selected and then a maximum number of trees per plot (different for each species) was further randomly subsampled.

Non-overlapping intervals were used to model diameter increments with models that lacked the path-invariance property, whereas mixed-length intervals were used for diameter difference equations and relative basal area models. Tree survival models were also constructed from both annual and mixed intervals.

A summary of the data coverage for each species is presented in Tables 10.1 to 10.3.

Table 10.1 Summary of individual tree data used for model development (*E. grandis*).

Variable	at	Non-overlapping intervals (n=5126)			Mixed-length intervals (n=5061)		
		Mean	Min.	Max.	Mean	Min.	Max.
age (years)	t_1	5.23	1.51	15.10	4.99	1.51	15.10
	t_2	6.23	2.48	15.83	6.26	2.48	15.83
dbh (cm)	t_1	16.5	1.4	42.5	15.7	1.2	44.0
	t_2	18.8	3.9	43.7	18.6	2.6	45.6
Relative basal area (R)	t_1	1.032	0.038	2.969	1.020	0.046	3.416
	t_2	1.023	0.043	2.598	1.003	0.093	3.051
Stand basal area (m^2/ha)	t_1	15.4	1.5	47.7	14.7	1.5	47.7
	t_2	19.6	3.0	49.6	20.1	2.6	49.6
Stocking (stems/ha)	t_1	804	200	1948	832	200	1948
	t_2	781	200	1875	801	200	1875

Table 10.2 Summary of individual tree data used for model development (*P. menziesii*).

Variable	at	Non-overlapping intervals (n=5073)			Mixed-length intervals (n=4447)		
		Mean	Min.	Max.	Mean	Min.	Max.
age (years)	t_1	33.21	10.85	81.05	28.80	7.25	75.00
	t_2	34.34	12.00	84.15	35.93	11.15	84.15
dbh (cm)	t_1	31.4	1.9	86.7	27.5	0.5	86.0
	t_2	32.1	2.1	89.6	31.5	1.3	88.1
Relative basal area (R)	t_1	1.050	0.032	5.574	1.063	0.005	5.487
	t_2	1.047	0.031	5.462	1.006	0.012	4.215
Stand basal area (m^2/ha)	t_1	42.1	3.8	109.5	40.7	3.6	105.9
	t_2	43.8	4.9	110.1	50.7	9.2	109.5
Stocking (stems/ha)	t_1	731	119	2850	914	49	2850
	t_2	724	119	2850	819	44	2850

A large number of explanatory variables including tree-level and stand-level variables, along with several transformations (logarithm, square root, square, inverse) were explored for the different tree-level model components (Table 10.4).

Table 10.3 Summary of individual tree data used for model development (*P. radiata*).

Variable	at	Non-overlapping intervals (n=5414)			Mixed-length intervals (n=4508)		
		Mean	Min.	Max.	Mean	Min.	Max.
age (years)	t_1	16.49	3.05	61.00	17.33	3.05	62.05
	t_2	17.50	4.05	62.05	26.36	4.05	64.05
dbh (cm)	t_1	29.5	1.1	100.1	29.5	1.1	104.1
	t_2	31.2	2.3	101.6	38.7	2.8	116.7
Relative basal area (R)	t_1	1.027	0.047	3.449	1.038	0.025	4.339
	t_2	1.025	0.046	3.436	0.971	0.039	3.482
Stand basal area (m^2/ha)	t_1	28.1	0.2	117.4	33.6	0.1	117.4
	t_2	30.7	1.0	118.5	51.8	3.2	118.5
Stocking (stems/ha)	t_1	450	54	2767	612	30	4444
	t_2	445	54	2650	533	30	4444

Table 10.4 Listing of tree-level and stand-level variables explored for tree-level model components.

Stand-level variables		Tree-level variables	
MTH	Mean top height	dbh	Diameter at breast height
G	Stand net basal area	g	Tree basal area
N	Stocking	h	Tree height
MTD	Mean top diameter	h_MTH	Tree height to MTH ratio
SI	Site index	h_dbh	Tree height to dbh ratio
ALT	Altitude	h_g	Tree height to tree basal area ratio
nn	Starting stocking (see Section 9.4.3)	BAL	Basal area of all trees with diameter larger than or equal to the subject tree
t	Age	DPIT	Diameter potential index of a tree
tN	Age by stocking product	CHG_PD	Change in potential diameter
tG	Age by basal area product	RD1	Relative diameter: dbh/QMD
QMD	Quadratic mean diameter	RD2	Relative diameter: dbh/MTD
Mean_g	Mean tree basal area	RBA	Relative basal area: g/G
MAI_MTH	Mean annual increment in MTH	BAL_d	BAL per dbh unit
MAI_MTD	Mean annual increment in MTD	NRD1	N*RD1
RS	Relative spacing: 1000/(MTH*sqrt(N))	NRD2	N*RD2
Rel_Dens	Relative density: G / QMD^0.5	CI	Competition index: 1 - BAL/G
SDI	Stand density index: $1.0147[10^{\wedge}(\log_{10}(N)+1.605*\log_{10}(QMD)-2.25]$		
Xt_tt	Thinning index: (1-dt/db)/tt #		
Xa_tt	Thinning index: (da/db-1)/tt #		
Gt_Gb_tt	Thinning index: (Gt/Gb)/tt (see 9.4.2)		

A useful set of variables was derived from the concept of diameter potential index of a tree (DPIT), in line with studies by Shula (1997a, 1997b, 1997c). DPIT is a measure of a tree's potential diameter growth, analogous to site index in that it is calculated for a reference age. From the estimated DPIT and given actual and future ages, the potential

diameter of a tree (PDT) can be calculated. This intermediate variable (not used itself as an explanatory variable) allows calculating the potential change in *dbh* (CHG_PD) for any projection interval. A brief explanation on how DPIT, PDT and CHG_PD were calculated is presented in Appendix 3.

Diameter increment models were constructed from (i) multiple linear regressions, (ii) exponential regressions (the ones used for modelling stand volume in Section 9.6), and (iii) models built with growth equations multiplied by logistic modifiers (e.g. Murphy and Shelton 1996). The number of independent variables for the former two approaches was limited to a maximum of five. For the third approach, the number of independent variables for the logistic modifier was limited to a maximum of four (the equations explored for the growth component contained two and three parameters).

The best models of each type were then compared and the best of the three was selected for further comparisons with the diameter models constructed using the difference-equation approach. All these comparisons were made using trees that were not included in datasets used for estimating model coefficients (Figure 10.1).

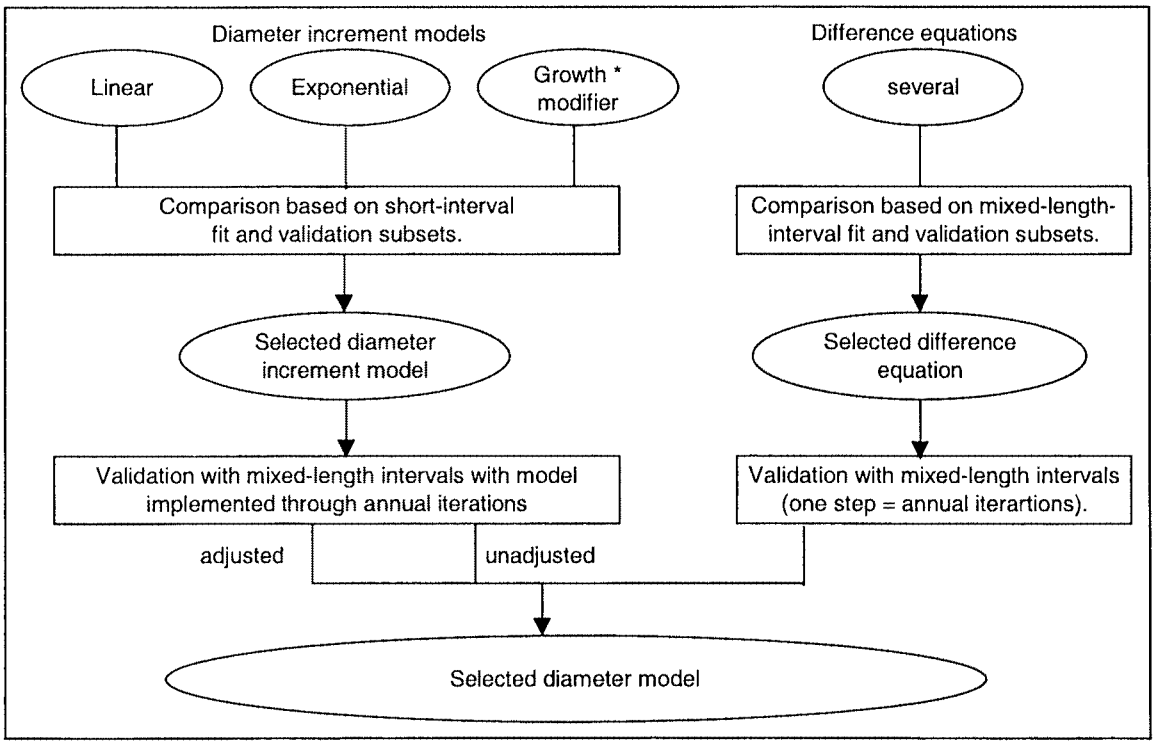


Figure 10.1 Flowchart describing the process for selecting the best diameter models.

As indicated in Figure 10.1, the validation of diameter increment models over longer projection intervals was implemented through annual iterations (i) by adjusting basal area and stocking estimates at each iteration through stand models, and (ii) without any stand-level-based adjustment. Adjustment procedures are described in Section 10.2.1. Before commencing these projections, each tree was assigned an expansion factor that indicated the number of trees that each tree represented in one hectare. The expansion factor was calculated from the plot area. For example, for a plot of 0.04 ha each tree in the plot represents 25 trees in one hectare (1/0.04), which is the same to say that each tree's expansion factor is 25. For both adjusted and unadjusted implementations, individual-tree survival models (see Section 10.3) were used to update expansion factors and compute stocking and other stand-density variables derived from stocking.

A number of sigmoid difference equations (those listed in Appendix 1 plus the model proposed by Zhang *et al.* 1996 shown in Section 7.7.3.1 as Equation 7.13) were fitted to diameter data. Models exhibiting the best fit were further improved by introducing site variables that were constant over time (e.g. site index, altitude). Introducing other stand variables such as changing basal area/ha would have removed the path-invariance property from the models.

Two relative basal area projection equations were tried. One was the original formulation proposed by Clutter and Allison (1974) and the other was an anamorphic difference equation derived from Schumacher's model. Zhao (1999) found that the latter model was almost as good as the former for *P. radiata* in Canterbury, New Zealand. The equations are, respectively:

$$R_2 = R_1^{(T_2/T_1)^{\beta}} \tag{10.1}$$

$$R_2 = R_1 \exp\left[\alpha\left(\frac{1}{T_1} - \frac{1}{T_2}\right)\right] \tag{10.2}$$

Logistic regressions for tree survival models were fitted using PROC LOGISTIC (SAS Institute Inc. 1989), which solves the parameters by maximum likelihood methods. The

same variables, transformations and interactions used for diameter increment models were explored through stepwise logistic regressions. Models for individual-tree annual survival were required to implement unadjusted individual tree models based on diameter increment equations and, therefore, needed to be developed first.

10.2.1 Dis-aggregative adjustments

For the adjusted implementation of diameter increment models, predicted tree diameters were modified so that the sum of all tree basal areas in a plot equalled the stand-level estimate of basal area. Likewise, the number of trees in a hectare represented by each tree (expansion factor) was adjusted so that summing the expansion factor of all trees in a plot equalled the stand-level estimate of stocking. Stocking adjustments were performed first, and adjusted expansion factors were used for adjusting tree diameters. Three approaches were explored for performing stocking adjustments, namely:

(i) Adjustment weighted by the reciprocal of diameter (Zhao 1999). With this approach, larger adjustments were made for smaller trees according to the following expression:

$$n^{*}_{2i} = \left(N_2 - \sum n_{2i}\right) \frac{n_{2i} dbh_{2i}^{-1}}{\sum n_{2i} dbh_{2i}^{-1}} \qquad (10.3)$$

where n_{2i} is the unadjusted expansion factor of the i^{th} tree; n^{*}_{2i} is the adjusted expansion factor and N_2 is the stand-level estimate of stocking.

(ii) Adjustment proportional to predicted expansion factor. With this approach, the original proportion of each tree's expansion factor in the implicit stocking was maintained, that is:

$$n^{*}_{2i} = \left(N_2 - \sum n_{2i}\right) \frac{n_{2i}}{\sum n_{2i}} \qquad (10.4)$$

(iii) Direct allocation of predicted stand mortality to each tree's expansion factor (Woollons and Hayward 1985). With this approach, the probability of tree mortality obtained from logistic models was used to calculate the mortality proportion to be allocated to each tree's expansion factor. Therefore, expansion factors were directly calculated by:

$$n_{2i}^{*} = n_{1i} - (N_1 - N_2)\frac{p_i}{\sum p_i} \tag{10.5}$$

where p_i is the probability of tree mortality (i.e. 1-P_{live}) of the i^{th} tree.

For the third method, predicted expansion factors after several annual iterations can be negative. When this occurred, the expansion factor was set to zero and the corresponding mortality yet to be allocated was distributed to the remaining trees in the PSP with a positive expansion factor. This procedure needed to be repeated a few times in some instances until the whole predicted stand mortality was successfully allocated. This mortality allocation approach was also used for projecting stand tables with the relative-basal-area-based approach (Chapter 12).

The stocking adjustment procedure that yielded the best diameter residual statistics for each species was selected. At each iteration, adjusted expansion factors were used to re-calculate predicted diameters using the following expression:

$$dbh_{2i}^{*} = dbh_{2i}\sqrt{\frac{G_2}{0.00007854\sum dbh_{2i}^{2} n_{2i}^{*}}} \tag{10.6}$$

Diameter adjustments obtained from Equation 10.6 corresponded to a constant percentage of increase or decrease of predicted diameter for all trees in a PSP.

10.3 MODELLING THE PROBABILITY OF TREE MORTALITY

Two different tree survival models were developed for each species. One model was required to predict the probability of annual tree survival for diameter increment models. The other model, which predicts the probability of tree survival for any projection interval, was required for the relative-basal-area-based dis-aggregative approach. Only non-overlapping intervals were used to fit the former model, whereas mixed-length intervals were used for the latter.

Tree- and stand-level variables shown in Table 10.4 plus some other interactions and transformations were considered for tree survival models. The variables kept in the models were selected through a stepwise regression approach using PROC LOGISTIC (SAS Institute Inc. 1989). The maximum likelihood method for estimating parameters, together with the fact of having a single observation per tree reduced the autocorrelation problem.

10.3.1 Probability of annual tree survival

Only intervals within the range of 0.8 to 1.2 years were used to fit these models. This constraint did not introduce any serious bias in terms of selecting greater proportions of measurements from younger plots. Selected variables with corresponding parameter estimates for the three species are displayed in Tables 10.5 to 10.7. The sign of the parameters indicates the type of effect on tree survival (i.e. positive ones indicate a positive effect on tree survival). The variables listed in Tables 10.5 to 10.7 are the X_i of the logistic model and corresponding parameter estimates are the a_i. The logistic model is given by Equation 10.7:

$$p=\frac{1}{1+\exp[-(a_0+a_1X_1+.....+a_nX_n)]} \qquad (10.7)$$

Table 10.5 Parameter estimates for the model of probability of annual survival (*E. grandis*).

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	12.7488	1.3994	82.9958	<.0001
QMD/dbh	-2.2526	0.1715	172.614	<.0001
1/SI	-252.8	51.0298	24.5492	<.0001
t/MTD	9.5064	2.6385	12.9812	0.0003
BAL^2/(dbh.1000)	-0.373	0.1312	8.0857	0.0045

Notes: QMD and MTD are, respectively, quadratic and top (i.e. largest 100 in a hectare) mean diameters of the stand in cm; BAL is the basal area (m^2/ha) of all trees with diameter larger than or equal to the subject tree; SI is site index and t is age (years).

All variables listed in Tables 10.5 to 10.7 are at the beginning of the period. The variable QMD/dbh is the inverse of relative diameter as defined by dbh/QMD. The negative sign of QMD/dbh implies a positive effect of relative diameter on survival, which was consistently found for all species. The variable t/MTD is the inverse of mean annual increment for MTD (MAI_{MTD}) and its positive sign implied a negative effect of MAI_{MTD} on tree survival.

Table 10.6 Parameter estimates for the model of probability of annual survival (*P. menziesii*).

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	-112.8	43.3917	6.7522	0.0094
MTD/t	53.5386	18.6704	8.2229	0.0041
QMD/dbh	-1.8688	0.3311	31.8665	<.0001
t/MTD	66.4709	25.1461	6.9875	0.0082

For Douglas-fir (*P. menziesii*), both MAI_{MTD} and its inverse were selected. The combined effect of these two terms was the same as for *E. grandis* (negative effect of MAI_{MTD} on tree survival), although the trend was curvilinear rather than linear.

Relative diameter and its exponential transformation were both selected for the annual tree survival model of *P. radiata*. The combined effect of relative diameter on tree survival was positive but the trend was curvilinear. Tree survival of *P. radiata* was

positively affected by relative spacing and negatively affected by mean annual increment in MTH.

Table 10.7 Parameter estimates for the model of probability of annual survival (*P. radiata*).

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	7.8053	1.5446	25.5346	<.0001
G/N	12.7613	4.1426	9.4895	0.0021
DPIT	0.1923	0.0602	10.223	0.0014
dbh/QMD	14.6006	2.8199	26.8086	<.0001
MTH/t	-1.8541	0.9134	4.1199	0.0424
$MTH*N^{0.5}$ / 1000	-2.933	1.3014	5.0796	0.0242
exp(dbh/MTD)	-9.1819	1.2351	55.2665	<.0001

Note: DPIT is the diameter potential index of the tree (see Shula 1997a, 1997b, 1997c and Appendix 3). The variable $MTH*N^{0.5}/1000$ is the inverse of relative spacing.

The performance of these models was evaluated with datasets formed with trees not used for model fitting. In the case of *P. radiata* these trees were from independent plots. Validation trees were grouped by diameter classes and predicted probabilities of tree survival were averaged. Predicted and actual probabilities of annual tree survival by diameter class are displayed in Figures 10.2 to 10.4. Predicted probabilities were very close to actual probabilities for most of the diameter range. In the smallest diameter classes there was more variability, as most of the mortality occurred in these classes.

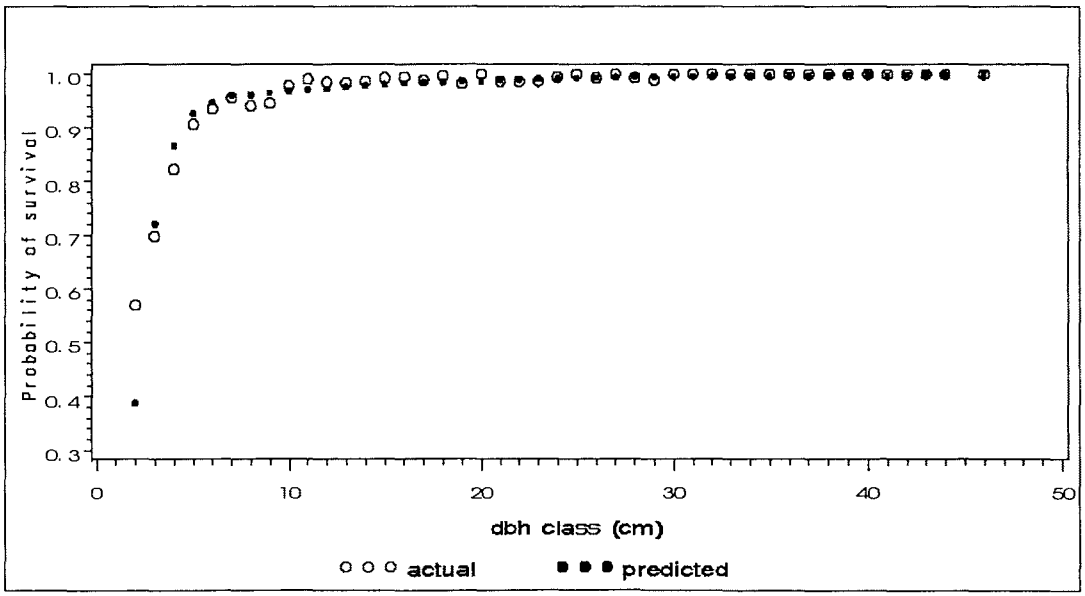


Figure 10.2 Actual and predicted probabilities of annual survival by *dbh* class (*E. grandis*)

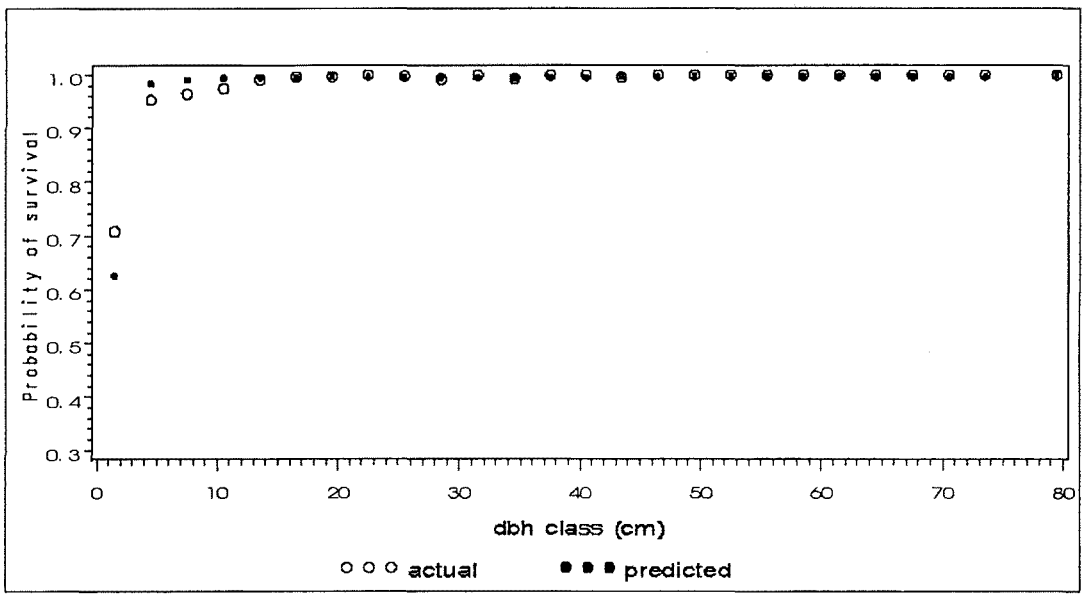


Figure 10.3 Actual and predicted probabilities of annual survival by *dbh* class (*P. menziesii*)

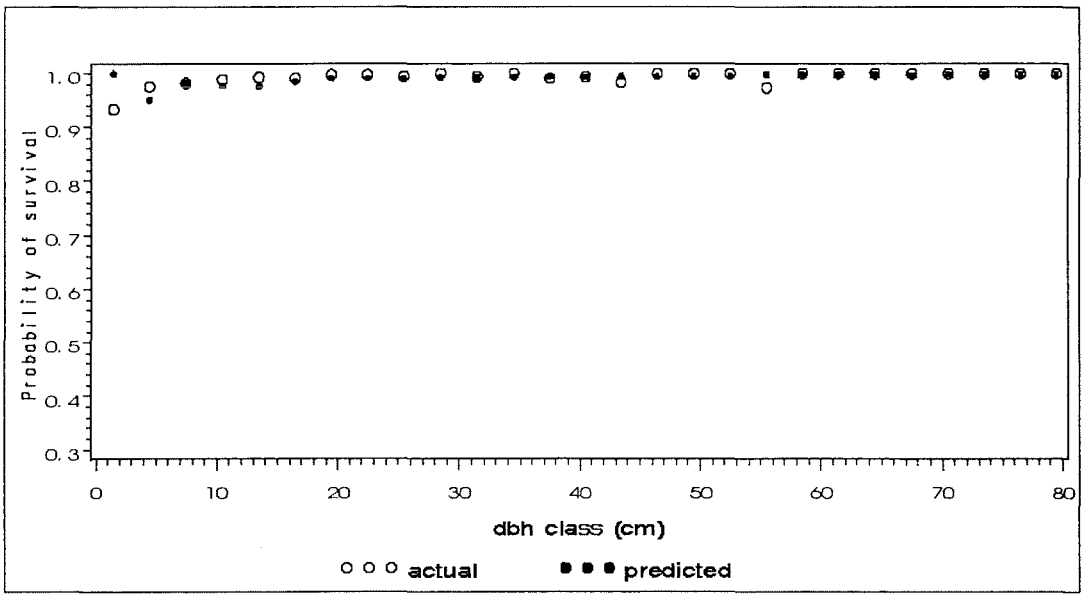


Figure 10.4 Actual and predicted probabilities of annual survival by *dbh* class (*P. radiata*)

10.3.2 Probability of tree survival for any projection interval

Mixed-length intervals were used to fit tree survival models for any projection interval. In order to account for the reduced probability of survival over longer projection intervals, the variable interval length was introduced and readily selected in the stepwise logistic regression procedure.

Selected variables with corresponding parameter estimates for the three species are displayed in Tables 10.8 to 10.10.

Table 10.8 Parameter estimates for the model of probability of tree survival for any projection interval (*E. grandis*)

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	-7.8414	0.9516	67.9014	<.0001
dbh/QMD	20.6119	1.8965	118.1258	<.0001
1/dbh	6.1133	1.2126	25.4154	<.0001
interval (years)	-0.4472	0.0741	36.4270	<.0001
R	-8.4065	0.9860	72.6921	<.0001

Note: R is relative basal area as previously defined.

All variables listed in Tables 10.8 to 10.10 are to be computed at the beginning of the interval. As expected, tree survival was negatively affected by the length of the projection interval for all three species. For *E. grandis*, the positive effect of relative diameter on tree survival was slightly counteracted by the negative sign of the parameter for R. The combined effect of relative tree size was positive but the trend was curvilinear.

Table 10.9 Parameter estimates for the model of probability of tree survival for any projection interval (*P. menziesii*).

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	4.0082	0.8609	21.6784	<.0001
interval (years)	-0.2111	0.0152	194.0824	<.0001
dbh/QMD	3.9777	0.7330	29.4495	<.0001
BAL/dbh	-0.6717	0.1039	41.7850	<.0001
CHG_PD	4.7182	0.8951	27.7854	<.0001
MTH/t	-3.9180	0.9212	18.0904	<.0001

Note: CHG_PD is the change in potential diameter (see Shula 1997a, 1997b, 1997c or Appendix 3). Other variables as previously defined.

For *P. menziesii*, mean annual increment in MTH negatively affected tree survival. The signs of all selected variables were consistent with their anticipated effects on tree survival.

Table 10.10 Parameter estimates for the model of probability of tree survival for any projection interval (*P. radiata*).

Parameter	Estimate	Std. error	Chi-Square	Pr > ChiSq
Intercept	4.5311	0.9464	22.9222	<.0001
interval (years)	-0.1425	0.0091	247.2943	<.0001
BAL/dbh	-1.4492	0.2102	47.5330	<.0001
BAL^2/dbh	0.0109	0.0022	24.4283	<.0001
1/dbh	11.6397	2.5749	20.4342	<.0001
dbh/MTD	16.5987	3.4270	23.4595	<.0001
exp(dbh/MTD)	-5.7424	1.5301	14.0840	0.0002

For *P. radiata*, BAL per unit *dbh* and relative diameter (as defined by *dbh*/MTD) showed curvilinear negative and positive effects on tree survival, respectively.

The above models were validated using trees that were not included in the datasets used for estimating model coefficients. In the case of *P. radiata* these trees were from independent plots. The same approach as for the models of annual probability of survival was used to ascertain the projection ability of these models. Plots in Figures 10.5 to 10.7 confirm the good performance of the models.

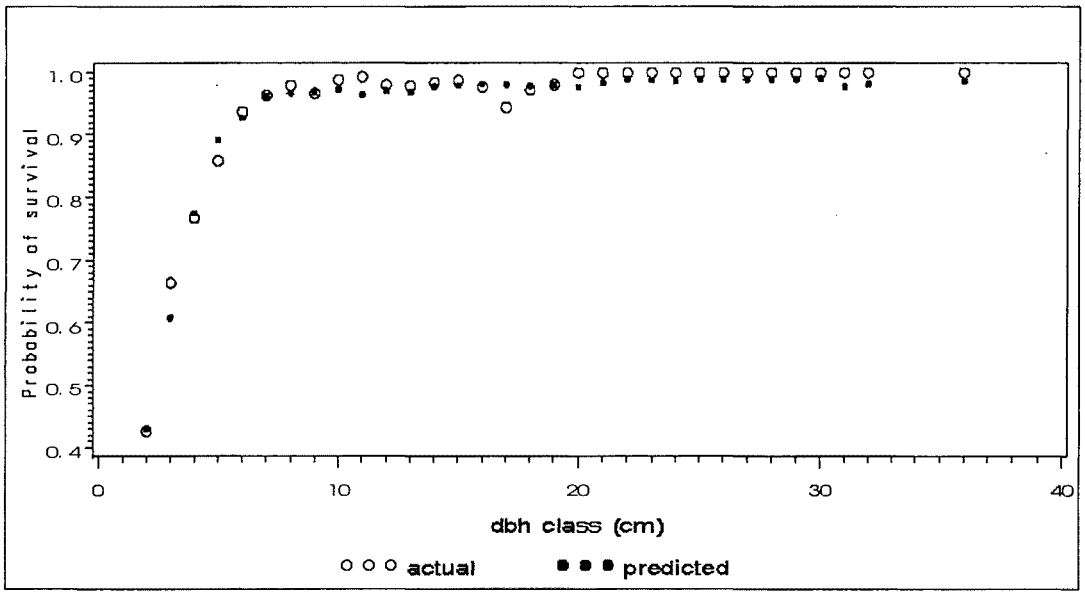


Figure 10.5 Actual and predicted probabilities of survival by *dbh* class (*E. grandis*)

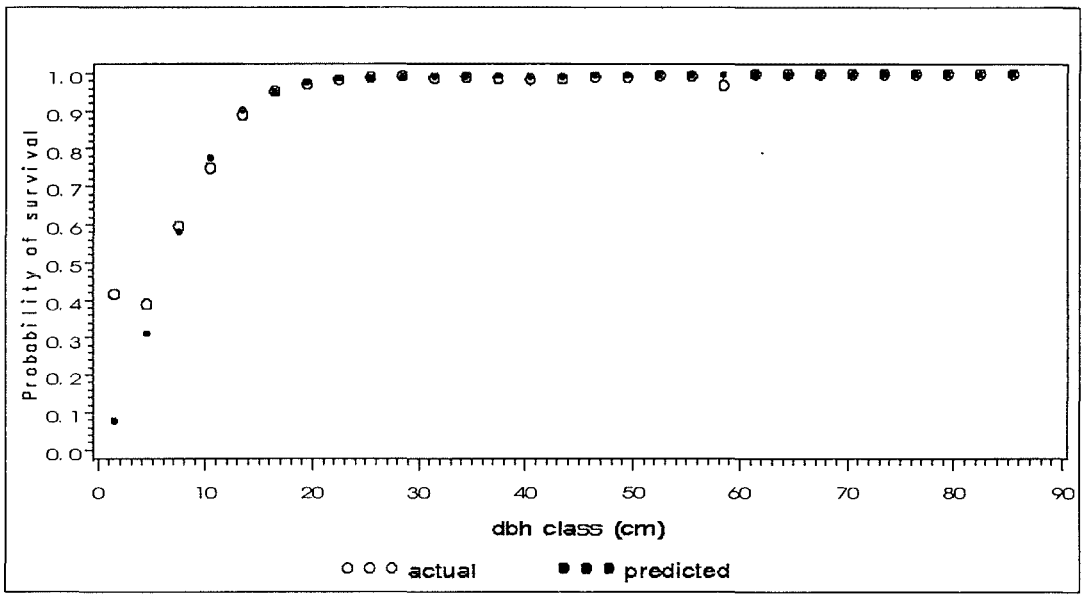


Figure 10.6 Actual and predicted probabilities of survival by *dbh* class (*P. menziesii*).

The high probability of survival at the smallest diameter class in Figure 10.7 was obtained from a validation plot measured at age 4.40 years. In this plot, despite the trees being small (low *dbh*), the probability of survival was close to 100% because of the small value of stand density measures such as BAL.

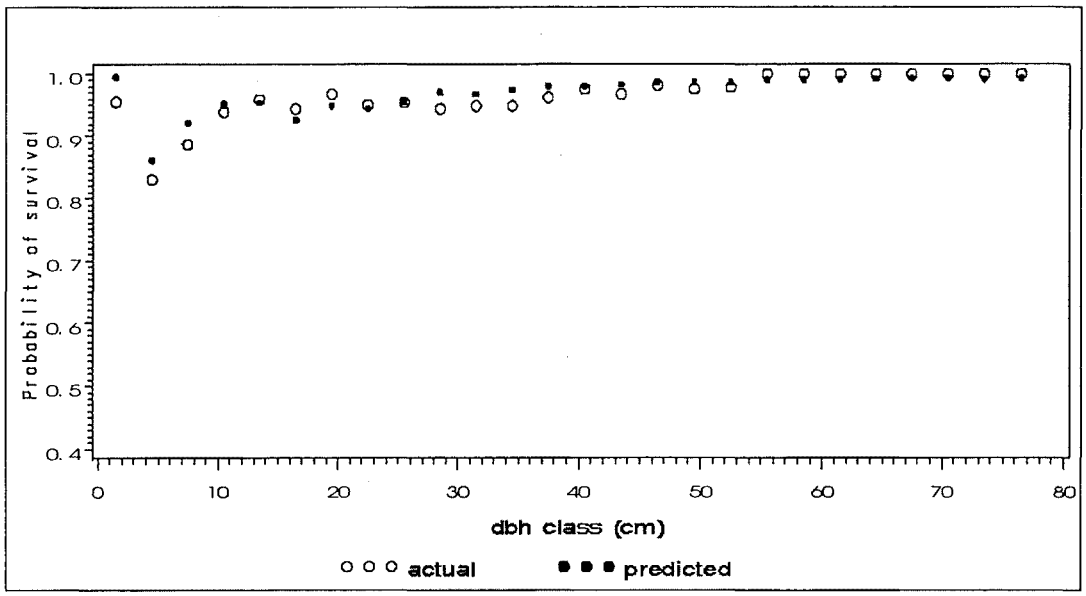


Figure 10.7 Actual and predicted probabilities of survival by *dbh* class (*P. radiata*)

10.4 MODELLING TREE DIAMETER GROWTH

10.4.1 Diameter increment models

Of the three approaches tried for modelling diameter increment, the simple multiple linear regression and the growth/modifier approach exhibited the best performance for *E. grandis*. Although residual statistics (Table 10.11) for these two models were relatively similar, after extensive residual plotting (not shown here) the linear model emerged as a more consistent model.

Table 10.11 Residual statistics of predicted diameter from diameter increment models (*E. grandis*).

	Fit subset (n=5126)			Validation subset (n=10264)		
Statistic	linear	exponential	G*M	linear	exponential	G*M
Mean	-0.015	-0.028	-0.021	-0.126	-0.164	-0.111
Std	0.86	0.88	0.88	1.17	1.20	1.20
Skewness	0.09	-0.01	0.04	-0.33	-0.39	-0.29
Min (1%)	-2.43	-2.62	-2.66	-3.65	-3.89	-3.76
Max (99%)	2.44	2.54	2.56	3.17	3.20	3.31

Note G*M is an approach based on a growth equation multiplied by a logistic modifier (Murphy and Shelton 1996).

Variables included in the selected (linear) model and parameter estimates are displayed in Table 10.12. When using this model, it is possible (although very unlikely) to get negative predicted diameter increments. To overcome this problem, predicted annual diameter increments need to be constrained to a minimum value. The minimum positive annual diameter increment in the dataset was 0.1 cm (a few trees having zero increment were recorded dead in the following measurement) and, therefore, this value will be used as the lower limit of predicted annual diameter increment for *E. grandis*.

Table 10.12 Parameter estimates and standard errors of the selected diameter increment model for *E. grandis* ($\Delta D=a_0 + a_1X_1 + ... + a_nX_n$).

Variable	Parameter Estimate	Std. Error	t value	Pr > | t |
Intercept	12.09742	0.61544	19.66	<.0001
dbh/QMD	1.96943	0.06390	30.82	<.0001
MTD/t	0.24611	0.01821	13.52	<.0001
ln(SDI)	-1.87719	0.07163	-26.21	<.0001
1/SDI	-77.66152	12.90548	-6.02	<.0001
1/SI	-46.79767	5.88252	-7.96	<.0001

Note: $SDI=1.0147*10^{\log_{10}(N_1)+1.605*\log_{10}(QMD)-2.25}$ (stand density index, Shula 1997b)

An exponential regression was the preferred approach to diameter increment modelling for *P. menziesii*. This equation produced the lowest mean residual for both fit and validation subsets (Table 10.13) and its residuals showed no untoward patterns when plotted against predicted values and independent variables.

Table 10.13 Residual statistics of predicted diameter from diameter increment models (*P. menziesii*).

	Fit subset (n=5073)			Validation subset (n=6392)		
Statistic	linear	exponential	G*M	linear	exponential	G*M
Mean	0.005	0.000	0.007	-0.008	-0.007	0.014
Std	0.34	0.34	0.34	0.29	0.28	0.28
Skewness	0.57	0.52	0.59	0.80	0.99	0.78
Min (1%)	-0.78	-0.81	-0.77	-0.64	-0.60	-0.64
Max (99%)	1.04	1.02	1.05	0.91	0.92	0.92

The exponential model, which parameter estimates are shown in Table 10.14, included the following variables:

$$\Delta D=\exp\left[a_0 + a_1\ln(\text{CHG\_PD}) + a_2\frac{\text{dbh}}{\text{QMD}} + a_3\frac{N}{1000} + a_4 RS + a_5 RS^2\right] \quad (10.8)$$

where RS (relative spacing) = $\frac{1000}{\text{MTH}\sqrt{N}}$, and other variables are as previously defined.

The signs of the parameters of all variables were consistent with their anticipated effects on tree diameter growth. The net effect of relative spacing was positive although the trend was curvilinear (second degree polynomial).

Table 10.14 Parameter estimates and standard errors of selected diameter increment models for *P. menziesii* and *P. radiata*.

	P. menziesii		*P. radiata*	
Parameter	Estimate	Std. error	Estimate	Std. error
a_0 / b_1	-1.2685	0.0883	4.09006	0.5162
a_1 / b_2	0.7306	0.0306	0.18302	0.0282
a_2 / x_0	0.3546	0.0378	1.06823	0.0979
a_3 / x_1	-0.1368	0.0140	-1.08775	0.0834
a_4 / x_2	0.5355	0.0321	9.43209	1.8976
a_5 / x_3	-0.0543	0.00421	0.05015	0.0024
/ x_4	-	-	0.36382	0.0312

For *P. radiata*, the best approach (as judged by residual statistics shown in Table 10.15 and residual plots) was to fit a growth function for ΔD on initial diameter multiplied by a logistic modifier limited to the interval [0-1]. This approach has been used in several recent individual tree models in the United States of America (Murphy and Shelton 1996; Bitoki *et al.* 1998; Lynch *et al.* 1999; Huebschmann *et al.* 2000). The selected growth function was a Chapman-Richards equation constrained by maximum size (Shifley and Brand 1984). The logistic modifier included CHG_PD, age, stand density index (SDI, Reineke 1933), and the basal area of all trees larger or equal than the subject tree (BAL) expressed by unit of *dbh*. The selected diameter increment model for *P. radiata*, the parameter estimates for which are shown in Table 10.14, was:

$$\Delta D = \frac{b_1 dbh^{b_2} - b_1 dbh/120(1-b_2)}{1+\exp(x_0 + x_1 CHG\_PD + x_2/t + x_3\sqrt{SDI} + x_4 BAL/dbh)} \quad (10.9)$$

Table 10.15 Residual statistics for predicted diameter from diameter increment models (*P. radiata*).

	Fit subset (n=5414)			Validation subset (n=4371)		
Statistic	linear	exponential	G*M	linear	exponential	G*M
Mean	-0.005	-0.007	-0.006	0.014	0.012	0.010
Std	0.64	0.64	0.63	0.71	0.70	0.70
Skewness	0.33	0.32	0.28	0.59	0.53	0.52
Min (1%)	-1.51	-1.57	-1.52	-1.72	-1.77	-1.77
Max (99%)	1.80	1.82	1.75	1.91	1.94	1.95

10.4.2 Adjusted *vs.* unadjusted implementation

Two implementation approaches were compared for the selected model of each species, namely an unadjusted implementation and an adjusted implementation. Adjustments were made with stand-level models of basal area and stocking at each iteration, as explained in Section 10.2.1. For the New Zealand datasets, the direct allocation of predicted stand mortality (Woollons and Hayward 1985) was the preferred adjusting approach. For *E. grandis*, adjustments weighted by the reciprocal of predicted diameter (Zhao 1999) yielded slightly better residual statistics of diameters. Basal area/ha and stocking/ha for the unadjusted implementation were derived from diameter increment and individual-tree survival models. Datasets used for these comparisons needed to contain all trees in each plot. A single plot interval was used for *E. grandis* and *P. menziesii*, whereas a maximum of two intervals per plot was used for *P. radiata*. The number of plot intervals was 272, 220 and 164 for *E. grandis*, *P. menziesii* and *P. radiata*, respectively, with projection intervals up to 6, 26 and 21 years, respectively.

For *E. grandis* and *P. radiata*, the overall accuracy (mean residual) of diameter models improved when stand-level-based adjustments were introduced (Table 10.16). However, the mean residual of the adjusted implementation was larger than the mean residual from the unadjusted implementation for *P. menziesii*. The overall precision, as judged by the

standard deviation of residuals, was unchanged by the implementation approach for all three species. For *E. grandis* and *P. menziesii*, the skewness of residuals decreased and positive and negative extreme residuals were more balanced with the adjusted implementation.

Residual statistics for annual projections for the three species (see Tables 10.11, 10.13 and 10.14) were clearly better than the statistics shown in Table 10.16, suggesting that errors were compounded over long projections.

Table 10.16 Residual statistics of predicted diameter from adjusted and unadjusted model implementations.

	E. grandis (n=18092)		*P. menziesii* (n=12236)		*P. radiata* (n=4374)	
Statistic	unadjusted	adjusted[1]	unadjusted	adjusted[2]	unadjusted	adjusted[2]
Mean	-0.09	0.03	-0.05	-0.32	-0.19	0.17
Std. dev.	1.10	1.10	2.04	2.04	2.88	2.88
Skewness	-0.55	0.39	0.94	0.15	-0.23	0.31
Min (1%)	-3.54	-2.94	-5.54	-6.87	-9.26	-8.27
Max (99%)	2.67	3.45	6.99	6.10	8.89	9.88

Notes: [1] expansion factor adjustments weighted by the reciprocal of predicted diameter; [2] direct allocation of predicted stand mortality (see Section 10.2.1)

The standard deviation of residuals (a measure of precision) for all three species were identical for both implementation approaches. For *E. grandis* and *P. radiata*, the adjusted implementation generated lower mean residuals than the unadjusted implementation (Table 10.16), and hence the former implementation approach was preferred for these species. However, the unadjusted implementation produced a lower mean residual for *P. menziesii*. There may be disadvantages of the unadjusted implementation when the whole modelling system, rather than diameters only, is considered. This aspect is addressed in Chapter 12.

10.4.3 Diameter difference equations

Both anamorphic and polymorphic difference equations were tried for diameter modelling. For the New Zealand datasets, anamorphic equations consistently outperformed their polymorphic counterparts. For the Uruguayan dataset, the relative

performance of anamorphic and polymorphic equations varied depending on models. The preferred base-model was the 2-parameter Schumacher polymorphic. The New Zealand datasets include a large proportion of trees from mature plots, whereas the Uruguayan dataset (*E. grandis*) is largely limited to young immature plantations. Perhaps, anamorphic equations would also seem to be better suited to diameter modelling of *E. grandis*, when data from older plots become available.

In order to maintain the intrinsic path-invariance property of classic difference equations, only variables that remained constant for each plot (e.g. altitude, site index) were considered for improving the fit of difference equations. Loss of the path invariance property might be a feasible option if the models were to be implemented through annual iterations.

Measures of stand density, a factor that is well known to play a key role in diameter growth, vary over time and, therefore, would not be appropriate for this approach. However, a simple measure of stand density that does remain constant over time is the nominal stocking at plot establishment or after thinning, that is, the stocking unaffected by mortality. This variable, which was termed "starting stocking" (to avoid confusion with initial stocking) and symbolised as *nn*, was incorporated in diameter projection equations.

The final diameter projection equation for *E. grandis,* along with corresponding parameter estimates, is shown in Table 10.17.

Schumacher anamorphic equations were preferred for the New Zealand datasets. The logarithm of starting stocking was successfully incorporated in these models indicating a non-linear effect of stocking on diameter growth. The logarithm of starting stocking modified different parameters of the Schumacher equation for both species (parameter *b* for *P. radiata* and parameter *c* for *P. menziesii*). Site index and altitude were also tried but they were either non-significant or produced only minor reductions in residual variances. Parameter estimates and standard errors of diameter projection equations for the New Zealand datasets are displayed in Table 10.18.

Table 10.17 Selected diameter projection equation for *E. grandis* (Equation 10.10).

$$dbh_2 = dbh_1^{\left(\frac{t_1}{t_2}\right)^b} \exp\left\{a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\}$$

$$a = a_0 + a_1 D_1 + a_2 D_2 + a_3 \frac{100}{nn} + a_4 SI$$

$$b = b_0 + b_1 D_1 + b_2 D_2 + b_3 \frac{100}{nn}$$

Parameter	Estimate	Srd. error
a_0	5.5589	0.1914
a_1	-0.6778	0.0852
a_2	-0.5730	0.0855
a_3	3.8241	0.2826
a_4	-0.0575	0.0054
b_0	0.4454	0.0219
b_1	0.4575	0.0445
b_2	0.2048	0.0329
b_3	-0.1329	0.0535

Where D_1=1 for Zone 8 (0 otherwise); D_2=1 for Zone 9 (0 otherwise); SI=site index and *nn* as defined in the text above.

Table 10.18 Selected diameter projection equations for *P. menziesii* (Equation 10.11) and *P. radiata* (Equation 10.12).

Species	Model	Parameter estimates (standard errors)			
		b/b_0	b_1	c/c_0	c_1
P. menziesii	$dbh_2 = dbh_1 \exp\left[-b\left(\frac{1}{t_2^c}-\frac{1}{t_1^c}\right)\right]$	13.2027	-	-0.1603	0.1582
	$c = c_0 + c_1 \ln(nn)$	(0.3545)	-	(0.0255)	(0.00319)
P. radiata	$dbh_2 = dbh_1 \exp\left\{-[b_0 + b_1 \ln(nn)]\left(\frac{1}{t_2^c}-\frac{1}{t_1^c}\right)\right\}$	16.9413	-1.6893	0.6392	-
		(0.4483)	(0.0632)	(0.0128)	-

Models were evaluated by analysing residual statistics (Table 10.19) and plots from developmental and validation subsets. For the New Zealand grown species, these datasets contained projection intervals of up to 30 years, whereas the Uruguayan dataset contained intervals of up to only 6 years.

Table 10.19 Residuals statistics from diameter projection equations for the three species

	E. grandis		*P. menziesii*		*P. radiata*	
Statistic	fit	validation	fit	validation	fit	validation
n	5061	18092	4447	12236	4508	4374
Mean	-0.07	-0.19	-0.12	0.05	0.17	1.24
Std. dev.	1.20	1.21	1.97	2.15	4.51	3.98
Skewness	0.25	0.08	0.42	1.03	0.17	1.30
Min (1%)	-3.39	-3.47	6.22	-5.74	-11.27	-8.40
Max (99%)	3.47	3.36	-5.38	7.65	12.91	15.90

For *E. grandis* and *P. menziesii* the models' mean residuals were very close to zero, despite the extended projection intervals for the latter. For the validation subset of *P. radiata*, on the other hand, the mean residual was large indicating an average underestimation of 1.24 cm. After analysing plots of residuals, it became readily apparent that most of the underestimation occurred for the youngest stands. The existing individual-tree diameter increment model for *P. radiata* in New Zealand (Shula 1997b) was built with data from PSPs aged 15 years or more (and its use correspondingly restricted), possibly because of the difficulty of modelling diameter growth at early stages. When the same restriction $(t_1 \geq 15)$ was imposed in this validation subset, the mean residual dropped to 0.24 cm, which represented a 0.6% of the average actual diameter (n=1875). In the case of imposing a restriction for young stands, MTH would be a more suitable index than age. When the validation subset was limited to initial mean top heights greater than or equal to 22 m (estimated MTH at age 15 for a site index of 30), the mean residual further decreased (0.16 cm, 0.4% of average dbh) despite the fact that a larger number of observations was kept (n=1958).

10.4.4 Diameter increment models *vs.* difference equations

The best diameter increment model for each species (i.e. adjusted implementation for *E. grandis* and *P. radiata* and unadjusted implementation for *P. menziesii*) was compared to selected diameter projection equations. This was performed through the examination of residual statistics computed from common datasets (Table 10.20).

Table 10.20 Diameter residual statistics from increment and difference equations

	E. grandis (n=18092)		*P. menziesii* (n=12236)		*P. radiata* (n=4374)	
Residuals' Statistic	ΔD model linear	Schumacher polymorphic	ΔD model exponential	Schumacher anamorphic	ΔD model G*M	Schumacher anamorphic
Mean	0.03	-0.19	-0.05	0.05	0.17	1.24
Std. dev.	1.10	1.21	2.04	2.15	2.88	3.98
Skewness	0.39	0.08	0.94	1.03	0.31	1.30
Min (1%)	-2.94	-3.47	-5.54	-5.74	-8.27	-8.40
Max (99%)	3.45	3.36	6.99	7.65	9.88	15.90

The magnitude of mean residuals for *P. menziesii* from the two methods was virtually identical despite the difference in sign. However, for *E. grandis* and *P. radiata* the diameter increment approach provided markedly lower mean residuals. The magnitude of the mean residual from the difference-equation approach for *P. radiata* was unacceptably large. As explained in Section 10.4.3, most of the underestimation of this model was for young stands.

The precision, as judged by the standard deviation of residuals, was greater with the diameter-increment approach for the three species. For the New Zealand datasets, positive and negative residuals were better balanced with the diameter-increment approach (as judged by the skewness and first and 99th percentiles). The opposite was the case for *E. grandis*.

Evidently, the diameter-increment approach was to be preferred for all three species.

10.5 MODELLING RELATIVE BASAL AREA

The sign of parameter β for the Clutter and Allison (1971) model (Equation 10.1) can be interpreted as the future contribution of individual trees in the projected stand. If β is positive, the relative contribution to stand basal area from trees larger than the average size (i.e. R>1) increases over time, whereas the relative contribution from trees smaller than the average size (R<1) decreases. Conversely, if β is negative, the relative contribution decreases for larger trees and increases for smaller trees.

A review of previous studies that used this model revealed that parameter β has been found to be both positive and negative (Table 10.21), even for the same species.

Table 10.21 Estimates of parameter β from previous studies. Model $R_2 = R_1^{(T_2/T_1)^\beta}$

Reference	Species	Estimated value of β
Clutter and Allison (1974)	*Pinus radiata* (radiata pine)	0.17
Pienaar and Harrison (1988)	*P. elliottii* (slash pine)	0.23
Pienaar (1989)	*P. elliottii* (slash pine)	0.05
Borders and Patterson (1990)	*P. taeda* (loblolly pine)	0.21
Knowe *et al*. (1997)	*Alnus rubra* (red alder)	-0.31
Zhao (1999)	*P. radiata* (radiata pine)	-0.16

The previous inconsistency regarding the trend of R suggested that the model $R_2=R_1$ (that is, R assumed constant over time) should also be tried along with the Clutter and Allison (1971) and Schumacher anamorphic models. Statistics of fit from these three models are displayed in Table 10.22.

Table 10.22 Mean square errors (MSE) and mean residuals (MR) for relative basal area models

	E. grandis		*P. menziesii*		*P. radiata*	
Model	MSE	MR	MSE	MR	MSE	MR
$R_2=R_1$	0.017	-0.014	0.029	-0.058	0.047	-0.067
Clutter and Allison (Eq. 10.1)	0.013	-0.004	0.023	-0.044	0.043	-0.054
Schumacher anam. (Eq. 10.2)	0.016	0.012	0.021	-0.001	0.042	-0.016

The model that assumed a constant relative basal area was clearly outperformed by the other two models. For the New Zealand datasets, the Schumacher model produced better statistics of fit than the Clutter and Allison (1974) model. However, the plots of residuals against predicted values were better for the latter model (Figure 10.8) and, therefore, the Clutter and Allison (1974) model was the selected base-model for the three species.

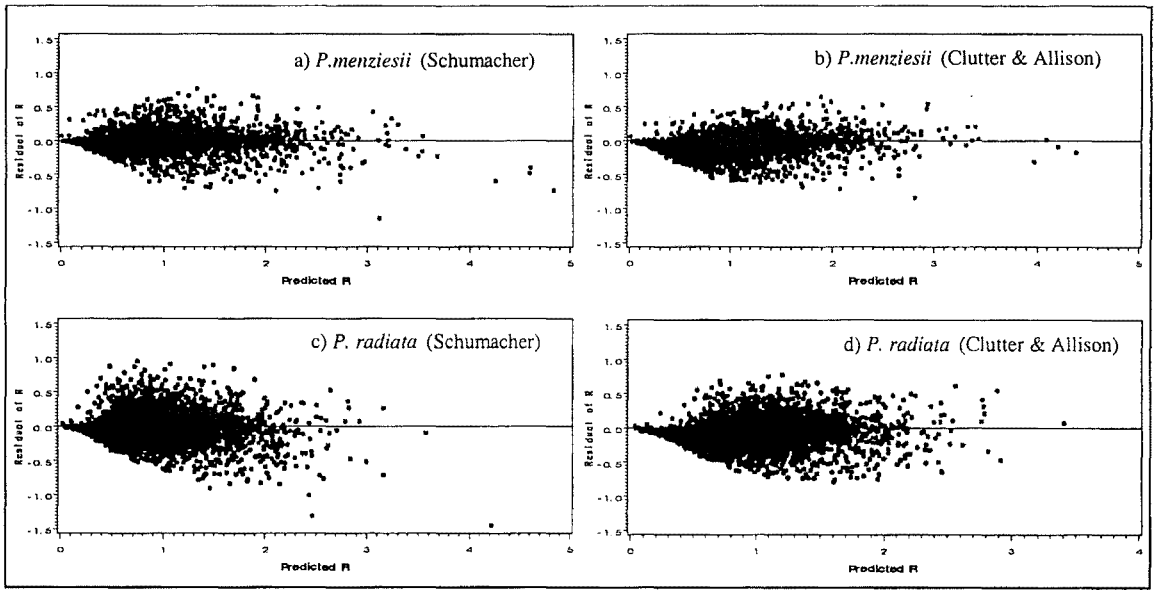


Figure 10.8 Plots of relative basal area (R) residuals against predicted values from two models, for *P. menziesii* and *P. radiata.*

Variations based on constants like site index, altitude and starting stocking were tried. For *E. grandis* and *P. radiata* no variable significantly reduced the residual variance. For *P. menziesii*, the inclusion of the variable starting stocking (nn) reduced the residual variance by 4.7%. However, the model that included *nn* performed worse than the base model for the validation dataset. Therefore, no modifications were made to the Clutter and Allison (1974) model for any species. Parameter estimates for the final relative basal area models are shown in Table 10.23.

Table 10.23 Parameter estimates and standard errors for parameter β (relative basal area models, Equation 10.1)

Species	Estimate	Std. error	Approximate 95% Confidence Limits	
E. grandis	-0.4351	0.0136	-0.4619	-0.4084
P. menziesii	-0.3180	0.0120	-0.3415	-0.2945
P. radiata	-0.2551	0.0157	-0.2859	-0.2243

For the three species, the value of β was negative indicating that the relative size of larger trees decrease and the relative size of the smaller trees increase over time. In other words, projected stands are more homogeneous than initial stands.

10.6 VERIFICATION AND VALIDATION

In addition to the overall validation presented in previous sections, diameter and relative-basal-area models were further evaluated through inspection of predicted-value and projection-interval classes[1]. This allowed a more thorough evaluation of the models, identifying situations where their performance seemed poorer than need be.

10.6.1 Diameter increment equations

For *E. grandis*, the selected diameter increment model (model in Table 10.12, adjusted) underestimated diameters for the diameter classes lower than 11 cm (Figure 10.9). For the rest of the diameter range error bars were largely confined to the interval [-1, 1]. The poorer performance of the model at the lowest *dbh* classes is unlikely to have noticeable impacts on stand statistics. However, this model should be re-fitted with an expanded dataset containing greater proportions of the smallest and largest trees.

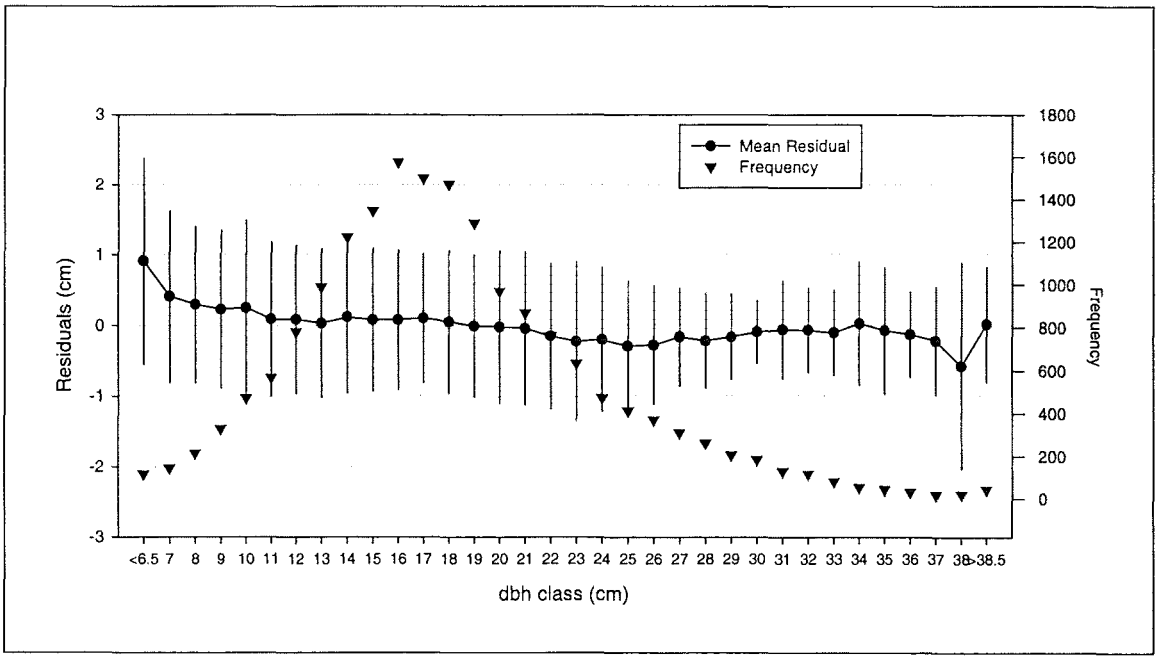


Figure 10.9 Mean residuals and standard errors (bars) of predicted diameters by diameter class (*E. grandis*).

[1] The limited range of projection intervals in the *E. grandis* dataset precluded rigorous analysis of this factor for this species

For *P. menziesii*, the diameter growth model selected as best (Equation 10.8, unadjusted) provided unbiased estimations for the whole range of predicted diameters and projection intervals (Figure 10.10). The oscillating pattern of mean residuals over the range of predicted diameters and projection intervals may be due to the particular incidence of certain plots at each point. The model's precision (as judged by the amplitude of standard error bars) deteriorated as the projection interval increased.

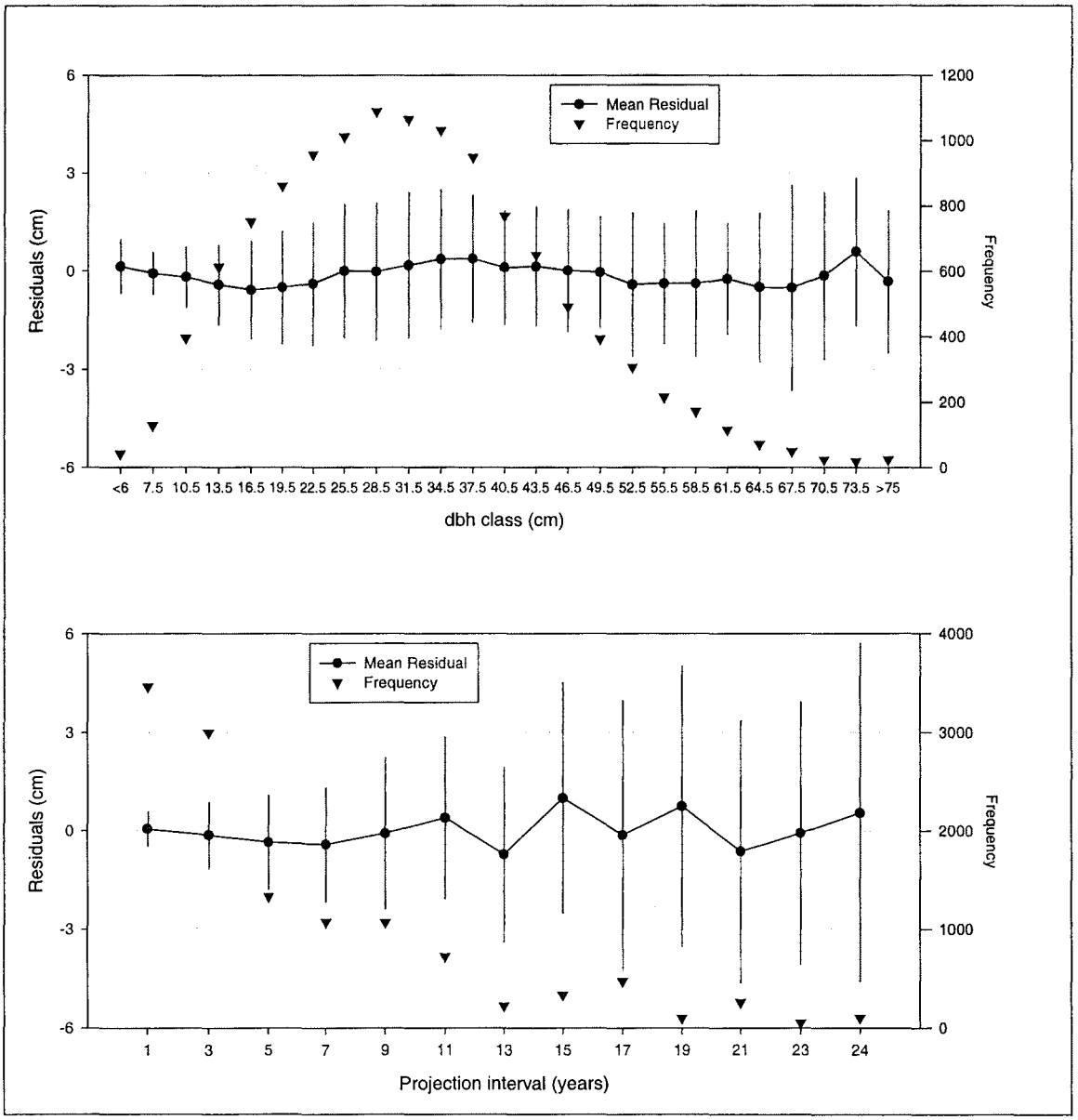


Figure 10.10 Mean residuals and standard errors (bars) of predicted diameters by diameter class (top) and interval class (bottom) - *P. menziesii*.

Mean residuals and standard errors of predicted diameters by diameter class and projection-interval class, for the *P. radiata* diameter increment model (Equation 10.9,

adjusted), are displayed in Figure 10.11. The model yielded unbiased projections over the whole range of predicted diameters. The graph at the bottom of Figure 10.11 shows that projections up to 4 years ahead were more accurate and precise than those for longer projections.

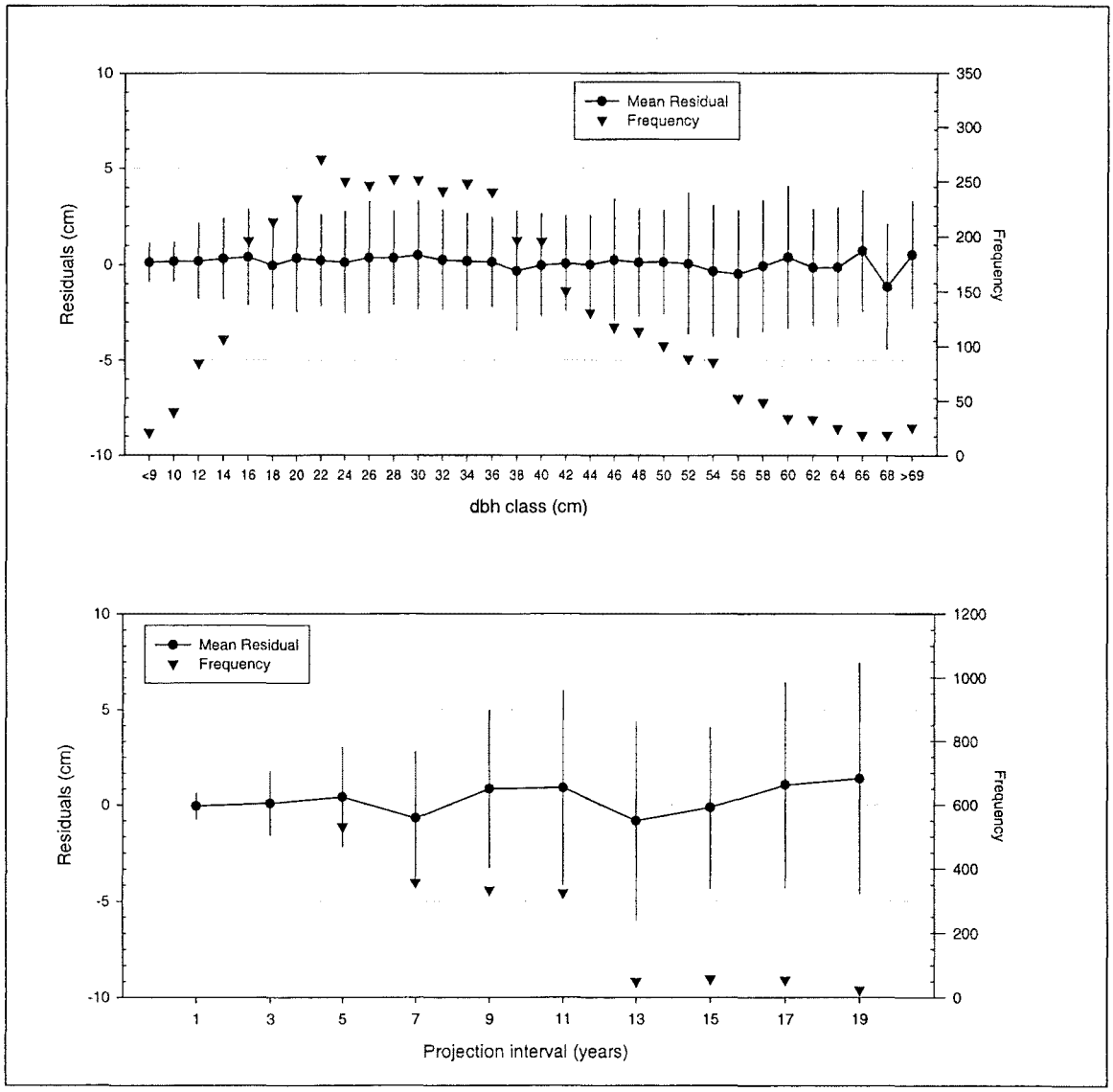


Figure 10.11 Mean residuals and standard errors (bars) of predicted diameters by diameter class (top) and interval class (bottom) - *P. radiata*

10.6.2 Relative basal area equations

For *E. grandis*, mean residuals of predicted R, using Equation 10.1 with β=-0.4351, exhibited a slight positive trend with predicted values (Figure 10.12). Relative basal areas were overestimated for trees with the smallest relative sizes and underestimated for trees

with the largest relative sizes. Given the negative sign of the parameter β, the previous trend indicates that the relative size of the smallest trees was projected to increase more than it actually did increase. Conversely, the relative size of the largest trees was projected to decrease more than it actually did decrease. These results suggest that the model was biased and that this component may need to be re-fitted to a dataset having a better coverage of relative tree sizes (in the event that the R-based dis-aggregative approach were to be adopted).

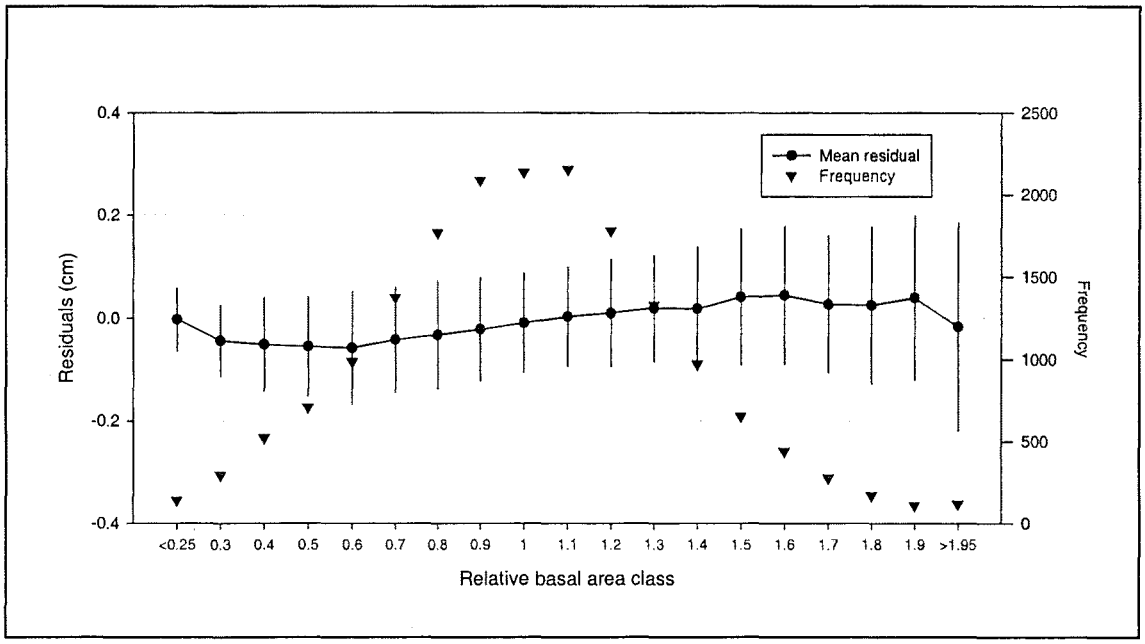


Figure 10.12 Mean residuals and standard errors (bars) of predicted relative basal area by R class (*E. grandis*).

Mean residuals for the relative basal area model for *P. menziesii*, namely Equation 10.1 with β=-0.318, showed an overall positive trend with predicted values (Figure 10.13), with a maximum bias of 0.126 at the 2.8 class (bias of 4.5%). The precision of projections decreased with increasing relative size and projection interval. The graph at the bottom of Figure 10.13 indicate that projections up to 5 years ahead performed acceptably in terms of accuracy and precision. Relative size was generally overestimated over longer projections. However, as predicted R would only be used as a weighting factor to apportion stand basal area to individual trees, the average overestimation over longer projections is of limited importance. On the other hand, the positive trend of residuals with R classes mentioned above would be of more practical relevance, as it would affect predicted diameter distributions by underestimating the diameter of the

largest trees and overestimating the diameter of the smallest trees. Overall, this model was imprecise and biased, perhaps because of the reduced number of observations for the large R classes in the dataset.

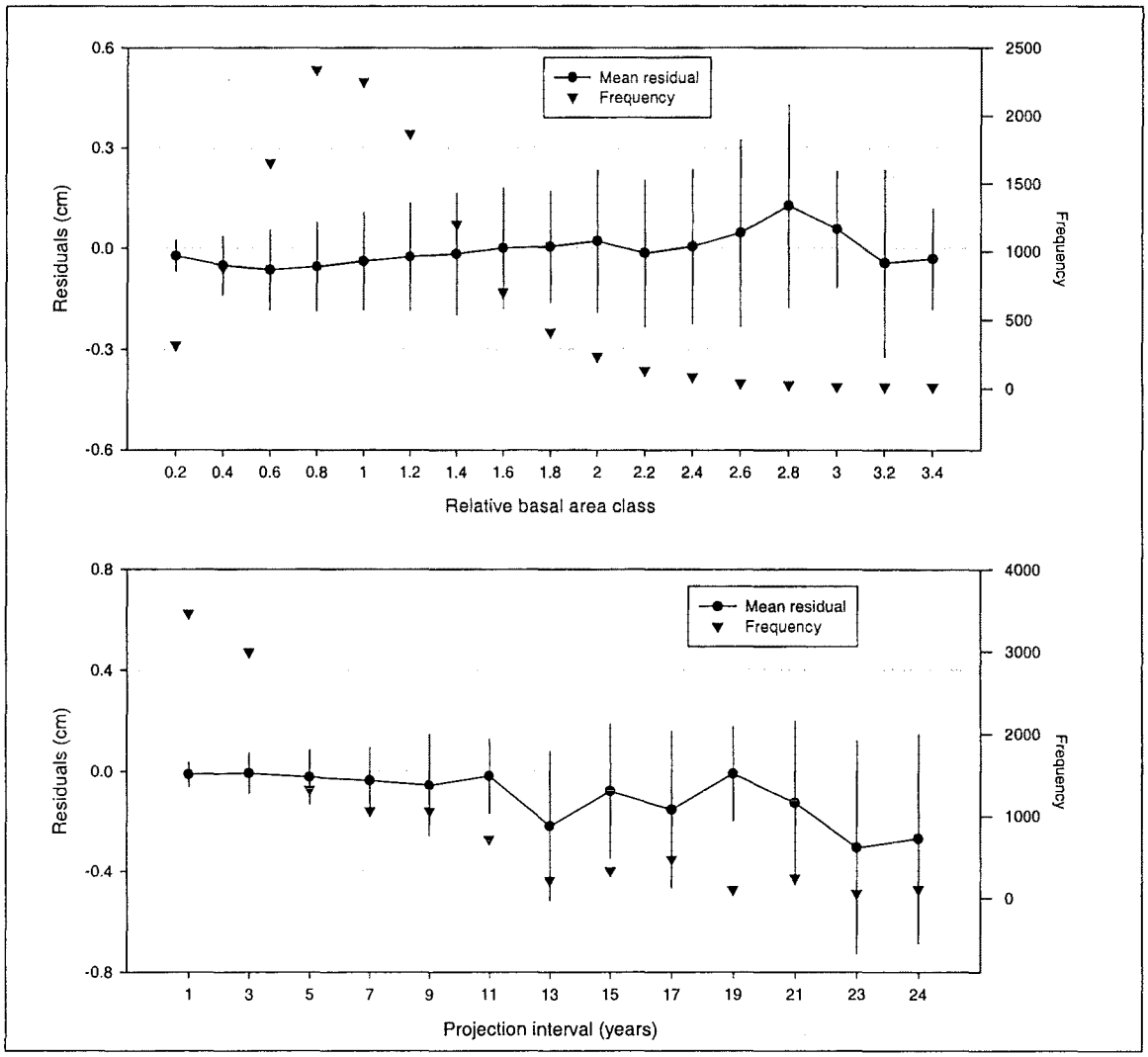


Figure 10.13 Mean residuals and standard errors (bars) of predicted relative basal area by R class (top) and interval class (bottom) - *P. menziesii.*

For *P. radiata*, the relative basal area model (Equation 10.1 with β=-0.2551) was accurate and precise for most of the range of predicted R (Figure 10.14). The accuracy was greatest between predicted R's of 1.0 to 1.2, slightly decreasing at both sides of this range. The model performed extremely well for projection intervals up to 17 years, with an expected reduction in precision as intervals increased.

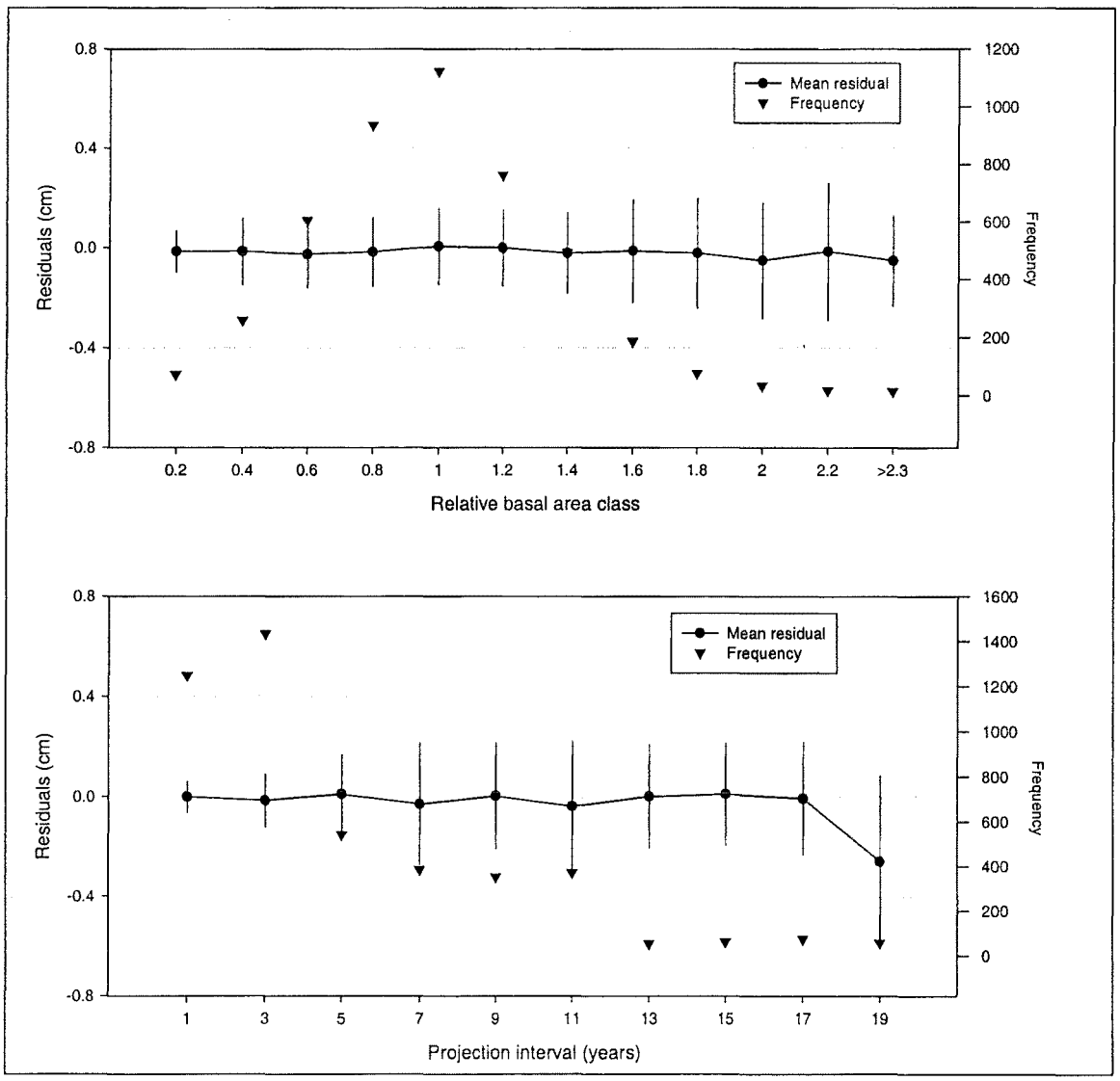


Figure 10.14 Mean residuals and standard errors (bars) of predicted relative basal area by R class (top) and interval class (bottom) - *P. radiata*

10.7 DISCUSSION AND CONCLUSIONS

The probability of tree survival was successfully modelled with the logistic model. Although this method was not compared to other form of model options, the results obtained from validating subsets and the logic of the sign and magnitude of the parameters for selected variables confirmed its reliability. Tree survival over long projection intervals was precisely modelled through including projection length as a predictor variable in the logistic model. Using this latter component in the relative-basal-area-based dis-aggregative approach will probably lead to better results than less specific assumptions concerning tree survival [e.g. Pienaar and Harrison (1988) assumed that tree

mortality was inversely proportional to R and Pienaar (1989) that it was inversely proportional to the square of R].

The relative merit of the three types of model evaluated for predicting diameter increment varied with species. A simple multiple linear regression was preferred for *E. grandis*, an exponential regression for *P. menziesii* and a growth-modifier approach for *P. radiata*. Exploring a broad range of explanatory variables including transformations and interactions was a key factor for achieving good projection capabilities for these models. For the New Zealand datasets, the variable potential change in diameter (or change in potential diameter, CHG_PD) made an important contribution to the selected models. This variable acted as a pre-estimate of growth itself based on the predicted pattern of dominant-tree diameter growth.

For *E. grandis* and *P. radiata*, diameter increment models performed better when annually adjusted by stand basal area and stocking models, than when they were implemented as stand-alone models. This may be explained by the compounding of errors from individual trees in the latter situation. However, for *P. menziesii*, the unadjusted diameter model produced better residual statistics than its adjusted counterpart. When deciding whether to use adjusted or unadjusted implementations for *P. menziesii*, compatibility and consistency issues may also have to be considered (see Chapter 12).

Diameter projection equations (difference equations) for *P. menziesii* and *E. grandis* produced only slightly higher means and standard deviations of residuals than diameter increment models did, . Conversely, the selected diameter difference equation for *P. radiata* was clearly biased (Table 10.20). Anamorphic models exhibited better fit than polymorphic ones for the New Zealand datasets, and Schumacher-based models were preferred for all three species. In this study, a decision was made to keep the path-invariance property of difference equations and, therefore, the scope for improving the base models was limited. Should this restriction be removed and the models implemented through adjusted annual iterations, the use of difference equations for diameter modelling may be appropriate. This topic may warrant further study.

The model proposed by Clutter and Allison (1974) for modelling the trend in relative basal area was confirmed as the best of those tried. The one-parameter Schumacher anamorphic model (Equation 10.2) exhibited better statistics of fit for the New Zealand datasets, but its residuals followed unacceptable patterns with predicted values. No additional explanatory variables were successfully introduced into relative basal area models.

The main conclusions drawn from the analysis of the performance of diameter-increment and relative basal area models by predicted-value and interval classes (Figures 10.9 to 10.14) are summarised in tabular form below (Table 10.24).

Table 10.24 Summarised qualitative description of the projection ability of ΔD and R models with validation datasets (from Figures 10.9 to 10.14).

Dataset	ΔD model	R model
E. grandis	negative trend of mean residuals with predicted values	positive trend of mean residuals with predicted values
P.menziesii	no obvious bias	trend of mean residuals with predicted value (+) and interval length (-)
P. radiata	no obvious bias	no obvious bias (slight overestimation for largest R classes)

For *E. grandis*, both model components exhibited poor results. This is attributed to limitations in the coverage and age of the dataset available. For the New Zealand grown species, diameter increment models compared favourable against relative basal area models. However, a formal comparison between methods needs to be carried out before drawing definitive conclusions. In such comparisons, the relative merit of the two approaches must be ascertained not only from a horizontal dimension (i.e. basal area and stocking) but also including the vertical dimension (height) so that volumes can be calculated. Consequently, the development of individual-tree height models is described in Chapter 11, whereas the final comparison of modelling approaches (including diameter distribution models) is discussed in Chapter 12.

CHAPTER 11

MODELLING INDIVIDUAL-TREE HEIGHTS

11.1 INTRODUCTION

As discussed in the Literature Review (Section 7.7.3.3), the most common approach to predicting individual-tree heights makes use of the relationship between tree diameter and tree height. This is because height growth is usually difficult to measure and predict given the magnitude of errors commonly associated with height measurements. Furthermore, tree heights are usually measured on just a small proportion of the trees in a plot and seldom is it possible to have height measurements of the whole plot population. This may preclude some modelling strategies.

Different methods have been followed to develop tree-height prediction systems based on height-dbh relationships. For those methods based on equations that predict height as a function of diameter (e.g. Petterson equation), the selection of the base model is crucial. Models with breast height (1.30 or 1.40 m) forced through zero *dbh* are the most appropriate as they attain an asymptote for large values of diameter and always have a positive slope (Curtis 1967). Linearity is another valuable property of height-dbh equations, particularly for PSP databases, as it allows the parameters to be explicitly and uniquely solved. For instance, the Petterson equation displayed in Section 7.7.3.3

(Equation 7.18), which is the standard height-dbh model in New Zealand, can be easily transformed into the following linear form:

$$Y=\frac{dbh}{(h-BH)^{0.4}}=\alpha dbh+\beta \qquad (11.1)$$

where Y is the dependent variable for the linear model, BH is breast height and α and β are the parameters to be estimated.

Another desirable feature of height-dbh equations is a reduced number of parameters. A single parameter would be ideal, although the prediction capability of such models might be excessively compromised. Zhao (1999) compared a number of two- and three-parameter equations and concluded that three-parameter equations offered limited benefits at large computational expense.

Thus, only two-parameter equations with an intercept equal to breast height were tried here.

Two methods for developing individual-tree height prediction systems, based on height-dbh equations, were explored and are discussed in following sections. A third method based on relating the ratio of height/MTH to the ratio of dbh/MTD (Gordon 1996; Shula 1997c) was also initially considered, but was clearly outperformed from early stages by the other two methods and therefore is not further discussed here.

Individual-tree height models developed in this chapter complement the individual tree models discussed in Chapter 10, and can be used to estimate the average height of each diameter class for the diameter distribution models described in Chapter 9.

11.2 METHODS AND PROCEDURES

The following two approaches were intensively analysed and compared for each species:

(i) The approach adopted in Zhao (1999), where the parameters of a model of the form h=f(dbh, α, β) are regressed on stand variables, and then the original model is refitted by substituting α and β with their corresponding sub-models.

(ii) The approach proposed by Lawrence (1990), where a model predicting the asymptote parameter (α) of the height-dbh equation is developed and the other parameter (β) is solved by substituting h and dbh with MTH and MTD, respectively.

For the first approach (referred to here as the *parameter prediction* approach) three sets of data were needed for each species, namely:

(a) a dataset with height-dbh pairs from all plot measurements, to screen for the most appropriate equation forms;

(b) a dataset with plot statistics and parameter estimates of the selected models, to fit parameter prediction sub-models; and

(c) a dataset with a subsample of height-diameter pairs, to estimate the final set of parameters.

The second dataset was formed with all plot measurements of *E. grandis* and *P. menziesii*, and a maximum of four randomly selected measurements per plot for *P. radiata*.

For Lawrence's approach (referred to here as the *asymptote prediction* approach) only the first two datasets were needed.

Both approaches were compared against validation datasets formed with tree measurements that were not included in the datasets used for fitting models through the

parameter prediction approach, that is dataset type (c) above. For *P. radiata*, the validation dataset contained trees only from the validation plots. The data coverage for the datasets used for parameter estimation and for model validation of the three species is shown in Tables 11.1 to 11.3.

Table 11.1 Summary of data used for developing and validating tree-height models of *E. grandis*

	Developmental subset (n=5254)			Validation subset (n=43827)		
Variable	Mean	Minimum	Maximum	Mean	Minimum	Maximum
t (years)	5.5	0.9	15.8	1.6	0.9	15.8
h (m)	16.5	1.2	46.6	5.5	0.7	47.3
dbh (cm)	15.6	0.8	47.6	5.7	0.4	46.2
N (stems/ha)	991	200	1948	336	200	1948
G (m^2/ha)	17.4	0.1	49.6	7.5	0.1	49.6
MTH (m)	18.3	3.0	45.1	5.2	2.8	45.1
MTD (cm)	19.5	2.3	42.1	5.5	2.3	42.1

Table 11.2 Summary of data used for developing and validating tree-height models of *P. menziesii*

	Developmental subset (n=3486)			Validation subset (n=21578)		
Variable	Mean	Minimum	Maximum	Mean	Minimum	Maximum
t (years)	32.7	7.3	83.1	32.1	7.3	84.2
h (m)	23.9	1.4	50.7	24.0	2.3	51.4
dbh (cm)	31.5	0.5	96.1	32.3	1.5	93.1
N (stems/ha)	759	44	2850	663	49	2850
G (m^2/ha)	40.2	3.6	110.1	39.9	3.6	110.1
MTH (m)	25.9	5.8	47.5	25.7	5.8	47.5
MTD (cm)	39.9	12.3	77.0	39.9	12.3	77.0

Table 11.3 Summary of data used for developing and validating tree-height models of *P. radiata*

	Developmental subset (n=5610)			Validation subset (n=12252)		
Variable	Mean	Minimum	Maximum	Mean	Minimum	Maximum
t (years)	21.4	2.2	64.1	16.2	4.3	43.1
h (m)	26.9	1.2	59.1	22.4	1.5	54.8
dbh (cm)	35.5	0.1	101.3	31.1	1.5	82.9
N (stems/ha)	523	30	4444	420	119	3877
G (m^2/ha)	37.7	0.1	118.5	27.8	0.4	83.5
MTH (m)	28.5	1.6	56.0	23.5	3.4	50.7
MTD (cm)	42.1	2.2	95.5	36.2	6.1	72.1

Seventeen 2-parameter equations having an intercept set to breast height and listed in Table 11.4, were evaluated and compared.

Table 11.4 Listing of the height-diameter equations assayed.

No.	Equation form	Example reference
11.3	$h=BH+\alpha dbh/(\beta+dbh)$	Huang *et al.* 1992, 2000
11.4	$h=BH+dbh^2/(\alpha+\beta dbh)^2$	Huang *et al.* 1992, 2000
11.5	$h=BH+(\alpha+\beta/dbh)^{-2.5}$	McEwen (1979)
11.6	$h=BH+(\alpha+\beta/dbh)^{-5}$	as above but exponent=-5. Zhao (1999)
11.7	$h=BH+(\alpha+\beta/dbh)^{-8}$	as above but exponent=-8. Zhao (1999)
11.8	$h=BH+\alpha(1-\exp(\beta dbh))$	Huang *et al.* 1992
11.9	$h=BH+\alpha(1+1/d)^{-\beta}$	Curtis 1967
11.10	$h=BH+\exp(\alpha+\beta/dbh)$	Schumacher 1939; Curtis 1967; Arabatzis and Burkhart (1992); Staudhammer and LeMay 2000
11.11	$h=BH+\alpha(\ln(1+dbh))^{\beta}$	Zhao (1999)
11.12	$h=BH+\exp(\alpha+\beta/(dbh+1))$	Huang *et al.* 2000
11.13	$h=BH+\alpha\exp(\beta/dbh)$	Huang *et al.* 2000
11.14	$h=BH+10^{\alpha}dbh^{\beta}$	Huang *et al.* 2000
11.15	$h=BH+\alpha dbh/(dbh+1)+\beta dbh$	Huang *et al.* 2000
11.16	$h=BH+\alpha(dbh/(1+dbh))^{\beta}$	Huang *et al.* 2000
11.17	$h=BH+\exp(\alpha)dbh^{\beta}$	Huang *et al.* 2000
11.18	$h=BH+\alpha dbh\ \exp(-\beta dbh)$	Huang *et al.* 2000
11.19	$h=BH+\alpha dbh+\beta dbh^2$	Huang *et al.* 2000

11.3 SCREENING FOR BEST EQUATION FORMS

For the first screening, candidate models (Table 11.4) were individually fitted to all plot measurements having at least 5 height-diameter pairs using PROC NLIN (SAS Institute Inc. 1989). Mean square errors (MSE) from each model were averaged to obtain an overall measure of model performance. In addition, the convergence percentage was also computed, as convergence failures were observed in some cases. The results of this procedure are presented in Table 11.5.

Table 11.5 Average MSE (MMSE), convergence percentage and ranking of height-diameter equations.

	E. grandis (n=929)			*P. menziesii* (n=2533)			*P. radiata* (n=8865)		
Equation	% conv.	MMSE	Rank	% conv.	MMSE	Rank	% conv.	MMSE	Rank
11.3	98.2%	0.8527	-	99.9%	2.0078	-	98.5%	2.3898	-
11.4	100.0%	0.8304	4	100.0%	2.0001	1	100.0%	2.3466	8
11.5	100.0%	0.8299	2	100.0%	2.0073	5	100.0%	2.3443	7
11.6	100.0%	0.8305	5	100.0%	2.0091	6	100.0%	2.3402	6
11.7	100.0%	0.8313	6	100.0%	2.0100	8	100.0%	2.3389	4
11.8	97.7%	1.4148	-	99.8%	2.0815	-	97.5%	2.4220	-
11.9	100.0%	0.8303	3	100.0%	2.0035	3	100.0%	2.3378	3
11.10	100.0%	0.8336	8	100.0%	2.0054	4	100.0%	2.3369	1
11.11	99.7%	0.8533	-	99.8%	2.0316	-	97.6%	2.3895	-
11.12	100.0%	0.8284	1	100.0%	2.0019	2	100.0%	2.3389	5
11.13	100.0%	0.8336	8	100.0%	2.0054	4	100.0%	2.3369	2
11.14	100.0%	0.8773	11	100.0%	2.0353	9	100.0%	2.4214	10
11.15	100.0%	0.9173	12	100.0%	2.0909	10	100.0%	2.4929	12
11.16	100.0%	0.8303	3	100.0%	2.0100	7	100.0%	2.3378	3
11.17	100.0%	0.8773	11	100.0%	2.0353	9	100.0%	2.4214	10
11.18	100.0%	0.8331	7	100.0%	2.1255	11	100.0%	2.3527	9
11.19	100.0%	0.8530	10	100.0%	2.4067	12	100.0%	2.4750	11

Note: Ranks were calculated only for equations having fully converged.

Equations 11.3, 11.8 and 11.11 failed to converge a few times for all species and, therefore, were not ranked. The Petterson equation with exponent -2.5 (Eq. 11.5) that is widely used in New Zealand ranked fifth and seventh for *P. menziesii* and *P. radiata*, respectively, although the differences in terms of average MSE (MMSE) within the group of top equations were minimal.

Given that a number of equations for each species produced virtually the same fitting statistics, the decision over which equation to select also needed to be based on the predictability of their parameters. For each species, the best eight equations were evaluated in that respect and four of them with high parameter predictability, as judged by the coefficient of determination (R^2), were selected as candidate base models (Table 11.6).

The coefficients of determination (R^2) displayed in Table 11.6 were considered just for comparison purposes and should not be regarded as a real measure of the proportion of variance explained by the models. This is because of the high level of autocorrelation of

the data from which they were calculated, which is well known to cause underestimation of error variances.

Table 11.6 Coefficients of determination (R^2) of multiple linear regressions fitted to parameters α and β of height-diameter equations

Independent Variables	*E. grandis* (n=932)			*P. menziesii* (n=2460)			*P. radiata* (n=3348)		
	Eq.	R^2 (α)	R^2 (β)	Eq.	R^2 (α)	R^2 (β)	Eq.	R^2 (α)	R^2 (β)
2	11.5	0.883	0.050	11.5	0.908	0.197	11.5	0.866	0.089
3		0.889	0.066		0.914	0.205		0.870	0.102
4		0.890	0.080		0.917	0.239		0.873	0.112
5		0.891	0.085		0.917	0.246		0.877	0.119
2	11.10	0.947	0.316	11.10	0.937	0.390	11.10	0.940	0.164
3		0.948	0.332		0.938	0.398		0.941	0.169
4		0.949	0.336		0.940	0.420		0.943	0.173
5		0.949	0.340		0.941	0.434		0.943	0.177
2	11.12	0.920	0.175	11.12	0.929	0.355	11.12	0.927	0.123
3		0.924	0.185		0.931	0.363		0.930	0.128
4		0.925	0.190		0.933	0.388		0.931	0.139
5		0.925	0.195		0.933	0.402		0.931	0.145
2	11.16	0.921	0.254	11.13	0.885	0.390	11.16	0.903	0.143
3		0.923	0.260		0.887	0.398		0.906	0.149
4		0.924	0.265		0.889	0.420		0.907	0.156
5		0.924	0.270		0.891	0.434		0.907	0.160

Note: Independent variables included the stand variables listed in Table 10.4 and transformations.

The Petterson equation with exponent -2.5 (Eq. 11.5) was outperformed by other equations in terms of parameter predictability. However, it was retained for further evaluation because of its popularity in New Zealand. For the three species, the log-reciprocal Schumacher equation (Eq. 11.10) exhibited the largest R^2 for both parameters with any given number of explanatory variables. Equation 11.12 was also selected for the three species and Equation 11.16 was selected for *E. grandis* and *P. radiata*.

The four base equations selected for each species were assayed for both the parameter prediction and asymptote prediction approaches, as explained in the following section.

11.4 DEVELOPING TREE HEIGHT MODELS

The R^2 values in Table 11.6 indicate that a great deal of the variation of the α parameter was explained by the first two explanatory variables and that the inclusion of additional independent variables only made minimal changes in the R^2. On the other hand, the R^2 values for the β parameter were lower and responded more to the addition of independent variables. Thus, two and five explanatory variables were initially assayed for parameters α and β, respectively, in the parameter prediction approach. Counting the intercept parameters of both sub-models (i.e. α_0 and β_0) the complete individual-tree height models from this approach would contain nine parameters.

Models with three explanatory variables (four parameters) were fitted to the α parameter for the asymptote prediction approach. Residual statistics for validation datasets are presented in Tables 11.7 to 11.9.

For *E. grandis,* estimations from the asymptote prediction approach were more precise (lower standard deviation of residuals) than estimations from the parameter prediction approach (Table 11.7). The accuracy of estimations (mean residual) exhibited the opposite trend (i.e. asymptote prediction approach less accurate than parameter prediction approach), except for Equation 11.12. For the parameter prediction approach, Equation 11.16 was clearly superior to the rest in terms of accuracy (mean residual), precision (standard deviation of residuals) and distribution of residuals (skewness). For the asymptote prediction approach, Equation 11.16 also exhibited the lowest mean residual, but the lowest standard deviation of residuals was achieved with Equation 11.12. Considering that the relative advantage of Eq. 11.16 in terms of accuracy was larger than the relative advantage of Eq. 11.12 in terms of precision, the former model was preferred.

Table 11.7 Residual statistics of predicted height from validation dataset (*E. grandis*, n=43827)

	Parameter prediction approach				Asymptote prediction approach			
Statistic	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq.11.16	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq.11.16
Mean	-0.014	0.024	0.041	0.000	0.037	-0.032	0.021	0.003
Std. dev.	0.899	0.919	0.903	0.896	0.885	0.894	0.881	0.888
Skewness	-0.165	-0.161	-0.244	-0.160	-0.163	-0.126	-0.155	-0.168
Min (1%)	-2.56	-2.56	-2.56	-2.52	-2.51	-2.54	-2.50	-2.54
Max (99%)	2.29	2.36	2.34	2.29	2.33	2.31	2.31	2.31
Parameters	9	9	9	9	4	4	4	4

For Douglas-fir (*P. menziesii*), some of the parameters selected for the linear sub-models of α and β were not significant when incorporated into the base equations. In fact, three parameters of Equation 11.10, two parameters of Equation 11.12, and one parameter of Equations 11.5 and 11.13 were not significant. Models were refitted without non-significant parameters, with the fit statistics remaining virtually unchanged. Altitude, which had been selected in the stepwise linear regressions for parameter β was removed in three of the four models. As for *E. grandis*, estimations from the parameter prediction approach were less precise but more accurate (except for Eq. 11.10) than estimations from the asymptote prediction approach (Table 11.8). For the parameter prediction approach, Equation 11.5 (Petterson equation) was deemed superior given that it exhibited the lowest standard deviation and skewness of residuals, and the second lowest mean residual. For the asymptote prediction approach, although Eq. 11.5 showed the lowest standard deviation and skewness, Eq. 11.13 generated the lowest mean residual and was therefore preferred.

Table 11.8 Residual statistics of predicted height from validation dataset (*P. menziesii*, n=21578)

	Parameter prediction approach				Asymptote prediction approach			
Statistic	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq. 11.13	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq. 11.13
Mean	0.050	0.089	0.083	0.045	0.125	0.084	0.114	0.074
Std. dev.	1.392	1.419	1.413	1.420	1.355	1.371	1.368	1.377
Skewness	-0.014	-0.117	-0.093	-0.053	0.056	0.073	0.064	0.125
Min (1%)	-3.32	-3.51	-3.47	-3.44	-3.17	-3.27	-3.24	-3.24
Max (99%)	3.39	3.52	3.53	3.54	3.57	3.59	3.59	3.63
Parameters	8	6	7	8	4	4	4	4

For *P. radiata*, all the equations fitted through the parameter prediction approach had one parameter that was not significantly different from zero. Therefore, the equations were refitted with eight parameters, with virtually no changes in error mean squares. As for the other two species, the asymptote prediction approach was more precise but less accurate (except for Eq. 11.12) than the parameter prediction approach. The selected equation for the parameter prediction approach was Eq. 11.16 given that it showed the lowest standard deviation and skewness and the second lowest mean residual. Equation 11.12 was clearly preferred for the asymptote prediction approach.

Table 11.9 Residual statistics of predicted height from validation dataset (*P. radiata*, n=12252)

	Parameter prediction approach				Asymptote prediction approach			
Statistic	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq. 11.16	Eq. 11.5	Eq. 11.10	Eq. 11.12	Eq. 11.16
Mean	-0.040	0.102	0.077	0.066	0.108	0.076	0.070	0.102
Std. dev.	1.486	1.381	1.357	1.333	1.329	1.313	1.312	1.312
Skewness	0.199	-0.165	-0.049	0.037	0.109	0.026	-0.004	0.066
Min (1%)	-3.41	-3.55	-3.38	-3.33	-3.26	-3.29	-3.30	-3.22
Max (99%)	3.78	3.51	3.56	3.56	3.65	3.51	3.45	3.57
Parameters	8	8	8	8	4	4	4	4

For none of the three species, was there a model obtained from either approach that clearly emerged as the best model. Within each approach, the selection of the best model was much easier. Therefore, a single model from each approach was further evaluated in more detail, by calculating mean residuals and standard errors by diameter class as illustrated in the following section.

11.5 FINAL COMPARISON, VERIFICATION AND VALIDATION

Data from validation datasets were stratified by diameter (dbh) classes and mean residuals and standard errors were calculated from the best model for each approach. This information is displayed in Figures 11.1 to 11.3 for *E. grandis*, *P. menziesii*, and *P. radiata*, respectively.

For *E. grandis*, mean residuals from the parameter prediction approach were consistently closer to the zero reference line (graph at the top of Figure 11.1) than mean residuals from the asymptote prediction approach (graph at the bottom of Figure 11.1). Equation 11.16 fitted through the parameter prediction approach was, therefore, the selected model for *E. grandis*. Residuals from the validation dataset computed from this model were very close to zero for most of the diameter range. The model underestimated tree heights of the lowest diameter class (2 cm) by 29 cm on average, suggesting that tree heights of the very smallest trees were not reliably estimated. Although the bias at the lowest dbh class was very high in relative terms, it is unlikely to cause noticeable effects on stand variables. The positive and negative mean residuals for the 36-cm and 38-cm classes, respectively, were probably due to the increased level of measurement errors for larger trees, plus the reduced number of observations available. Overall, the model performed satisfactorily with a global mean residual of 0.03 cm and 90% of the 43827 residuals between -1.5 and 1.5 m.

Parameter estimates for the selected model of *E. grandis*, namely Equation 11.16 fitted through the parameter prediction approach, are presented in Table 11.10.

As for *E. grandis*, mean residuals for *P. menziesii*, from the parameter prediction approach (graph at the top of Figure 11.2) were better aligned with the zero reference line than mean residuals from the asymptote prediction approach (graph at the bottom of Figure 11.2). Therefore, Equation 11.5 (Petterson equation) fitted through the parameter prediction approach was the preferred individual-tree height model for *P. menziesii*. The model underestimated heights for the smallest diameter classes, but presented residuals very close to zero throughout the 26-cm to the 65-cm classes, which contained the bulk of data. For the largest diameter classes, mean residuals oscillated between positive and negative values possibly because of the increased level of measurement errors and reduced number of observations in these classes. The model underestimated tree heights by 5 cm on average, but 90% of the 21578 residuals were between -2.2 and 2.4 m suggesting an acceptable overall performance

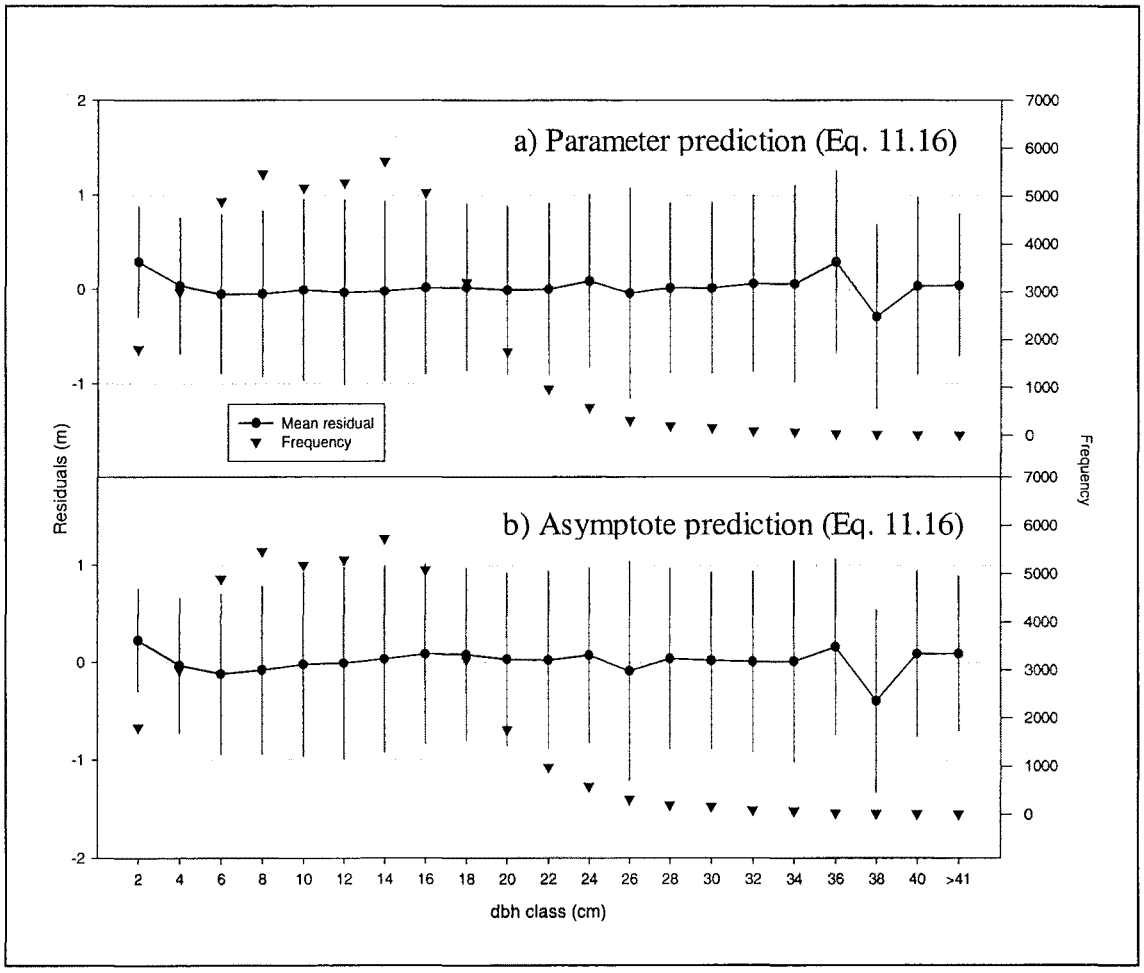


Figure 11.1 Mean residuals and standard errors of tree heights from validation dataset (n=43827), grouped by dbh classes (*E. grandis*)

Table 11.10 Parameter estimates for the selected model for *E. grandis* (Equation 11.16)

Parameter	Estimate	Std. error	Approximate 95% Confidence Limits	
a_0	3.6011	0.2516	3.1077	4.0944
a_1	1.4479	0.0172	1.4143	1.4816
a_2	-0.3114	0.0185	-0.3477	-0.2751
b_0	8.1424	0.6867	6.7960	9.4887
b_1	0.2379	0.0287	0.1815	0.2942
b_2	-0.7051	0.0854	-0.8726	-0.5376
b_3	-0.0345	0.0085	-0.0512	-0.0178
b_4	0.2782	0.0444	0.1912	0.3652
b_5	3.3316	0.3631	2.6197	4.0435
α	$a_0 + a_1MTH + a_2MTD$			
β	$b_0+b_1t+b_2Ln(N)+b_3G+b_4\frac{MTD}{t}+b_5\frac{MTH\sqrt{N}}{1000}$			

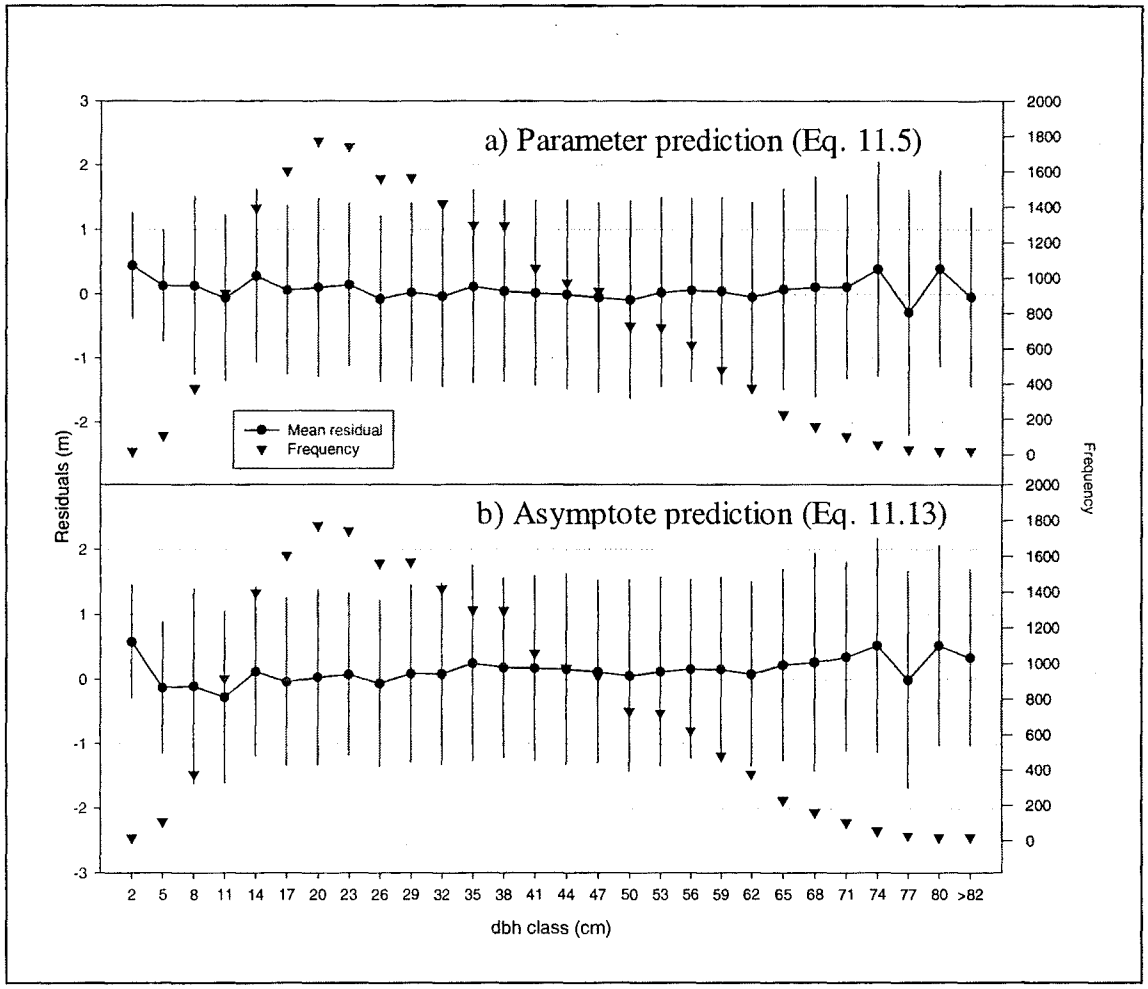


Figure 11.2 Mean residuals and standard errors of tree heights from validation dataset (n=21578), grouped by dbh classes (*P. menziesii*)

Parameter estimates for the selected model of *P. menziesii*, namely Equation 11.5 (Petterson equation) fitted through the parameter prediction approach, are presented in Table 11.11.

Unlike the other two species, mean residuals for *P. radiata* were better aligned with the zero reference line for the asymptote prediction approach (graph at the bottom of Figure 11.3) than for the parameter prediction approach (graph at the top of Figure 11.3). Consequently, Equation 11.12 fitted through the asymptote prediction approach was the selected model for *P. radiata*. Mean residuals were located slightly over the zero reference line for most of the diameter range, indicating that tree heights were slightly underestimated. The overall underestimation was 7 cm and 90% of the residuals were between -2.1 and 2.2 m. As for the previous two species, mean residuals of the largest

diameter classes oscillated between positive and negative values indicating the increased variability of tree heights in this area (perhaps caused by larger measurement errors).

Table 11.11 Parameter estimates for the selected model for *P. menziesii* (Equation 11.5)

Parameter	Estimate	Std. error	Approximate 95% Confidence Limits	
a_0	0.5429	0.0052	0.5327	0.5530
a_1	0.0079	0.0005	0.0070	0.0088
a_2	-0.0989	0.0013	-0.1015	-0.0964
b_0	0.4905	0.0837	0.3263	0.6547
b_1	0.1117	0.0344	0.0442	0.1792
b_2	0.1150	0.0122	0.0911	0.1389
b_3	0.3010	0.0776	0.1489	0.4530
b_4	0.2374	0.0110	0.2157	0.2590
α	$a_0 + a_1 Ln(G) + a_2 Ln(MTH)$			
β	$b_0 + b_1 \frac{ALT}{1000} + b_2 \frac{t^2}{1000} + b_3 \frac{MTH}{t} + b_4 Ln\left[\left(\frac{1000}{MTH\sqrt{N}}\right)^2\right]$			

Parameter estimates for predicting the asymptote parameter α of the selected height-diameter model for *P. radiata*, namely Equation 11.12, are presented in Table 11.12. Unlike the parameter prediction approach, where the parameters were estimated by non-linear regressions, the parameters for the asymptote prediction approach were estimated through linear regression.

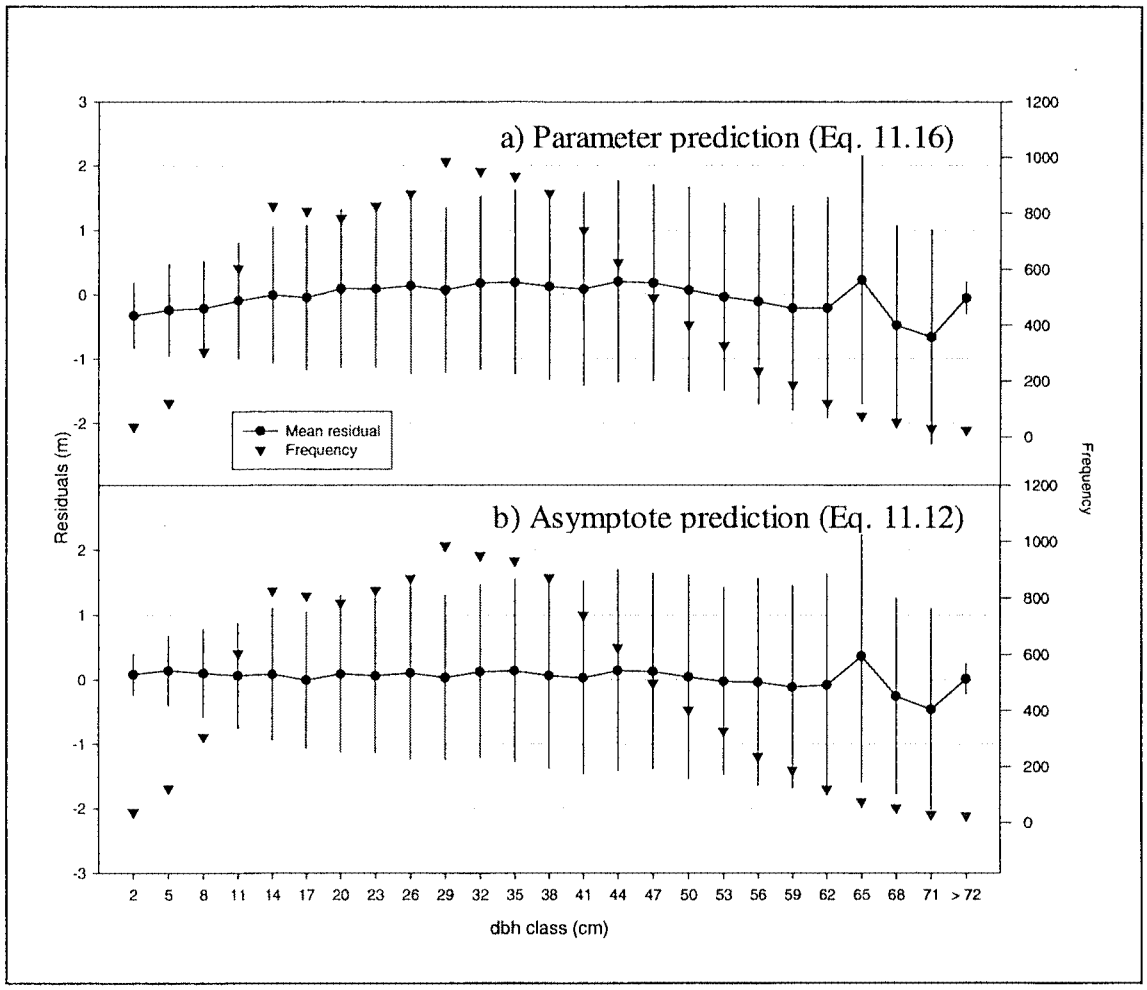


Figure 11.3 Mean residuals and standard errors of tree heights from validation dataset (n=12252), grouped by dbh classes (*P. radiata*)

Table 11.12 Parameter estimates for predicting the α parameter of Equation 11.12 (*P. radiata*).

Parameter	Estimate	Std. error	Type II SS	F Value	Pr > F
a_0	0.51156	0.0154	21.97	1104.8	<.0001
a_1	0.94221	0.0068	378.53	19032.8	<.0001
a_2	0.30489	0.0220	3.82	192.3	<.0001
a_3	-0.15990	0.0081	7.83	393.7	<.0001
α	$a_0+a_1\mathrm{Ln}(\mathrm{MTH})+a_2\frac{\mathrm{MTH}\sqrt{\mathrm{N}}}{1000}+a_3\mathrm{Ln}\left(\frac{\mathrm{G}}{\mathrm{QMD}^{0.5}}\right)$				

Note: QMD is the quadratic mean diameter, and other variables as previously defined

The parameter β was solved by substituting h and dbh in Equation 11.12 (Table 11.4) with MTH and MTD, respectively. Making β the subject in equation 11.12, results in:

$$\beta=[\mathrm{Ln}(\mathrm{MTH}-1.4)-\alpha](\mathrm{MTD}+1) \tag{11.20}$$

11.6 DISCUSSION AND CONCLUSIONS

From the list of candidate models displayed in Table 11.4, Equation 11.10 (Schumacher equation) was the one with the best parameter predictability (Table 11.6). However, this equation was not selected as the best model for any of the species or modelling approaches, as parameter predictability alone should not be a decisive criterion for selecting equation forms. Neither was the average mean square error (calculated by fitting the candidate equations to actual height-diameter pairs for each plot measurement) a critical indicator of how the equations behaved under both modelling approaches. With regard to average mean square error, the selected equations ranked third for *E. grandis* and fifth for the New Zealand grown species (Table 11.5).

Different equations were chosen for each species, which indicated that the selection of equation forms for modelling tree height was data dependent. Both modelling approaches were found appropriate, the parameter prediction approach being preferred for *E. grandis* and *P. menziesii* and the asymptote prediction approach for *P. radiata.*

Previous considerations indicated that the thorough and multistage process followed was justified to select the equation form and modelling approach best suited to each dataset.

The Petterson equation (Eq. 11.5), routinely used in New Zealand, was the equation form selected to model tree heights of *P. menziesii.* However, this equation produced very poor results for *P. radiata*, particularly when combined with the asymptote prediction approach (Table 11.9). The three selected equations have the desirable properties of passing through breast height at zero diameter, approaching an asymptote, having always a positive slope and being transformable to a linear form.

The performance of the final models was scrutinised by diameter classes. This procedure enabled a more powerful assessment of the models than was possible by analysing overall statistics of fit. It also allowed identification of situations where they exhibited poorer performance.

In general, the prediction capability of the models was acceptable both in terms of accuracy and precision. The model for *P. radiata*, which exhibited the highest overall bias, overestimated heights by 7 cm on average. For this species, the validation dataset was truly independent of the data used for model fitting (independent PSP's), which may explain the larger bias. Given the high level of measurement error for tree heights, this bias is likely to be within the error margin, particularly for larger trees.

Models developed in this chapter can be used to estimate individual-tree heights of all trees in a PSP from their diameters, provided some height measurements are taken to compute MTH. In the context of this study, these models were required as inputs to compute individual tree volumes for the comparison of growth modelling approaches discussed in Chapter 12.

CHAPTER 12

COMPARISON OF MODELLING APPROACHES

12.1 INTRODUCTION

Different modelling systems, all producing both stand-level and tree-level (or diameter class) output resolution have been developed and discussed in previous chapters. Diameter distribution models (Chapter 9) do not retain individual tree information but can provide an output resolution almost as detailed as individual tree models can, provided that the width of the diameter classes is set to a very small value.

Individual tree models were built through two general approaches, one based on modelling each tree's relative hierarchy within the stand over time (relative basal area) and the other based on modelling diameter increments on an annual basis (Chapter 10). The former is a dis-aggregative approach whereby projected stand statistics can be apportioned to individual trees. The diameter increment approach can be a stand-alone model completely independent of the stand model or it can be adjusted so that basal area and stocking estimates are compatible with the stand model. When individual-tree models were ascertained in terms of diameter estimation capabilities, the adjusted implementation produced more accurate estimations than the unadjusted implementation for *E. grandis* and *P. radiata.* However, the unadjusted implementation produced more

accurate estimations for *P. menziesii* (see Section 10.4.2). Nevertheless, comparisons of both implementation approaches for individual tree models need to consider the whole modelling system (rather than diameters alone). The main attribute of a stand from the forest managers' perspective is probably the volume, or more precisely the merchantable assorted volume. Tree volumes can be estimated from inputs of *h* and *dbh* into an appropriate two-dimensional tree volume equation. Individual-tree height models were developed in Chapter 11.

Comparisons of models obtained with different modelling approaches can be performed at two levels. Stand level comparisons are aimed at ascertaining the models' ability to estimate stand statistics such as basal area and stocking. The second level includes the assessment of diameter distributions (or stand tables) generated by the models. At both levels, model predictions need to be compared with actual plot measurements.

Although the evaluation of predicted stand tables is more complete than the stand-level appraisal (and somehow inclusive of it), the latter may provide useful insights into the weaknesses and strengths of the models. Both procedures were, therefore, adopted for this study. For comparing the prediction ability of diameter distributions, the error-index approach was followed (Reynolds *et al.* 1988; or see Section 7.7.4).

12.2 METHODS AND PROCEDURES

The same validation datasets used in Chapter 10 for testing diameter increment models and diameter difference equations over mixed projection intervals were used in this chapter. For *P. radiata*, the validation dataset contains up to 2 randomly selected intervals from independent plots (not used to fit any model component). For *E. grandis* and *P. menziesii*, the validation datasets contain a single randomly selected interval from each plot (for these species all plots were used for model development as explained in Chapter 8). Validation datasets contain all trees in each of the selected plots and projection intervals. The data coverage in datasets used for evaluating and comparing modelling approaches is summarised in Table 12.1.

Table 12.1 Summary of tree and stand variables in datasets used for comparing modelling approaches.

		E. grandis (n=18792)			*P. menziesii* (n=13296)			*P. radiata* (n=4653)		
Variable	at	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.	Max.
Plots	-		272			220			85	
Projection intervals	-		272			220			168	
Interval length (yr)	-	1.3	0.6	5.9	5.6	0.9	25.9	4.9	0.8	20.0
Age (yr)	t_1	5.3	3.1	15.1	29.6	7.3	74.0	13.6	4.4	35.0
	t_2	6.5	3.8	15.8	36.3	11.2	78.0	18.5	6.3	43.1
dbh (cm)	t_1	15.6	1.3	46.2	27.0	0.5	89.7	24.7	1.5	76.7
	t_2	18.0	3.6	47.6	31.6	2.1	92.2	32.8	3.7	82.9
R	t_1	1.00	0.03	3.33	1.00	0.00	5.48	1.00	0.05	4.10
	t_2	1.00	0.04	3.31	1.00	0.02	4.54	1.00	0.05	3.13
h (m)	t_1	16.3	2.7	44.6	21.6	1.4	47.4	18.0	1.6	45.4
	t_2	19.7	6.3	47.3	25.8	2.3	50.6	25.0	3.3	53.9
G (m^2/ha)	t_1	17.0	5.0	47.7	36.2	5.7	100.7	22.3	1.1	63.9
	t_2	21.6	8.5	49.6	44.5	9.5	106.2	33.7	3.3	75.1
N (stems/ha)	t_1	803	260	1643	609	49	2850	361	119	2260
	t_2	773	250	1507	560	44	2850	343	119	1780
MTH (m)	t_1	18.8	9.8	42.7	25.0	5.8	45.1	21.5	4.0	41.4
	t_2	22.6	13.3	45.1	28.8	9.0	45.5	28.4	6.4	49.3
Volume (m^3/ha)	t_1	153	21	836	364	16	1240	202	3	770
	t_2	220	59	911	491	42	1342	353	12	1150

Note: The 85 plots of *P. radiata* were not used to fit any model component and, therefore, constitute an independent validation dataset.

Basal area and stocking estimations from unadjusted individual tree models were conducted by annually updating tree diameters and expansion factors with diameter increment (Section 10.4.1) and tree survival models (Section 10.3.1), respectively.

For each species, stand tables were generated using the following four methods: (i) diameter distribution model; (ii) relative basal area dis-aggregative approach; (iii) unadjusted individual tree model; and (iv) adjusted individual tree model. Fields in each stand table included: (a) diameter class midpoint; (b) frequency (number of trees in a hectare contained in the diameter class); (c) height of the tree of midpoint diameter; (d) average tree volume of the diameter class.

For the New Zealand datasets, tree volumes were estimated from diameter class midpoints and predicted average tree heights of the diameter class using tree volume tables 10 and 136[1] for *P. radiata* and *P. menziesii* which are, respectively, given by:

$$TV=dbh^{1.711}\left(\frac{h^2}{h-1.4}\right)^{1.196}\exp(-10.18942) \tag{12.1}$$

$$TV=dbh^{1.828198}\left(\frac{h^2}{h-1.4}\right)^{1.102592}\exp(-10.19719) \tag{12.2}$$

For *E. grandis*, tree volumes were estimated with Equation 12.3, which was derived from a variant of the Max and Burkhart (1976) taper equation (Equation 2.11a with parameters in Table 5.2).

$$TV=\left\{K\left[\left(\frac{b_2}{3}\right)+\left(\frac{b_1}{2}\right)-(b_1+b_2)+\left(\frac{b_3\ln(dbh)}{3}\right)a_1^3+\left(\frac{b_4\ln(dbh)}{3}\right)a_2^3\right]\right\}dbh^2h \tag{12.3}$$

where K=0.00007854

Once the stand tables were obtained, error indices (Reynolds *et al.* 1988) were calculated as follows:

$$EI=\sum\left|(actual\_freq_{ij}-pred\_freq_{ij})W_{ij}\right| \tag{12.4}$$

where $actual\_freq_{ij}$ and $pred\_freq_{ij}$ are actual and predicted frequencies, respectively, of the i^{th} diameter class for the j^{th} plot projection interval and W_{ij} is the weighting factor.

Different weighting factors were explored, including tree volume, tree basal area and no weighting factor (i.e. a difference of one tree in the 15-cm diameter class treated equally as a difference of one tree in the 40-cm diameter class). Analysis of variance and

[1] Tree volume tables obtained from the software FFCALC Version 1.3 (Forest Research Institute 1992)

multiple range tests were used to determine whether differences in error indices (for a given weighting factor) were statistically significant.

Graphical methods were also used to assess and compare modelling approaches. These included scatter plots of mean error indices against projection interval and plots of predicted diameter distributions overlapped with histograms of actual distributions.

12.3 RESULTS

12.3.1 Stand-level comparisons

Residuals of stand basal area and stocking from unadjusted individual tree models (ITM) and from stand-level equations were calculated for each species. Residual statistics are displayed in Tables 12.2 (basal area) and 12.3 (stocking).

Table 12.2 Statistics of stand basal area residuals from unadjusted individual tree models (ITM) and stand models

Residual statistic	*E. grandis*		*P. menziesii*		*P. radiata*	
	ITM	Stand	ITM	Stand	ITM	Stand
Mean	0.35	-0.03	0.24	0.13	0.08	-0.05
Std. dev.	1.26	1.27	3.38	2.71	4.03	3.62
Skewness	0.28	0.52	0.69	0.99	-1.25	-0.55
Min. (1%)	-2.7	-3.3	-10.7	-6.4	-19.2	-14.0
Max. (99%)	4.0	3.9	12.5	8.3	10.8	10.5

For all datasets, stand-level equations for basal area outperformed individual tree models (ITM) for basal area/ha estimations in terms of accuracy (mean residual), and balance between positive and negative extreme residuals (Table 12.2). For *E. grandis*, the standard deviation of residuals from the ITM was slightly lower than the standard deviation of residuals from the stand-level equation. In addition, residuals from the *E. grandis* ITM showed a more uniform distribution around the zero reference line when plotted against predicted and initial basal areas (Figure 12.1) than residuals from the stand-level equation.

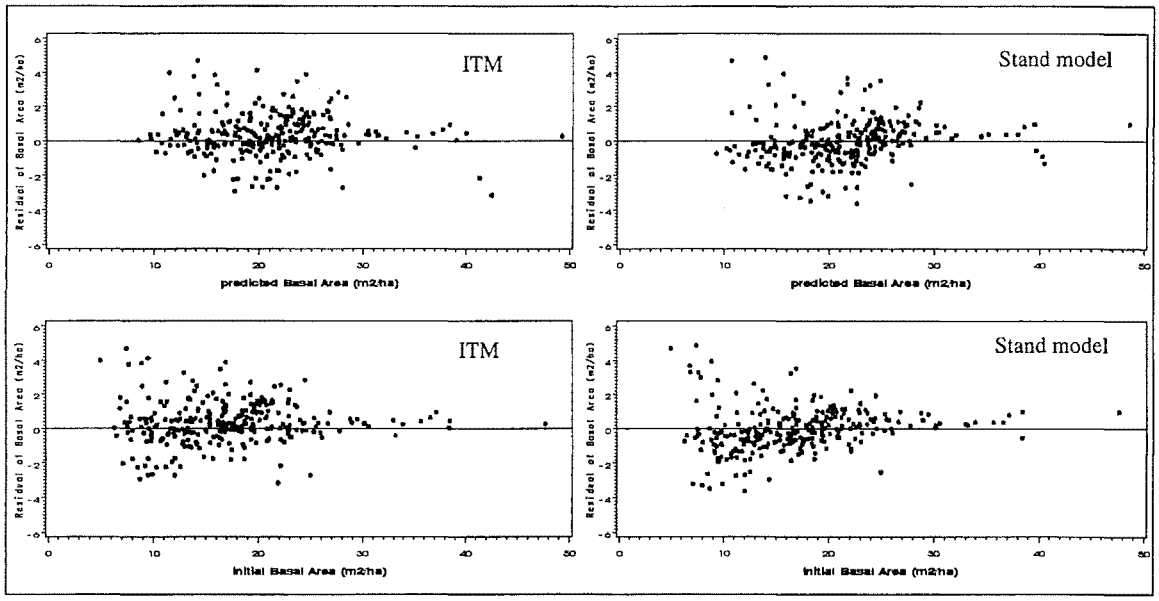


Figure 12.1 Basal area residuals (m^2/ha) against predicted values (top) and initial basal area (bottom) for the individual tree model (left) and stand level model (right) - *E. grandis.*

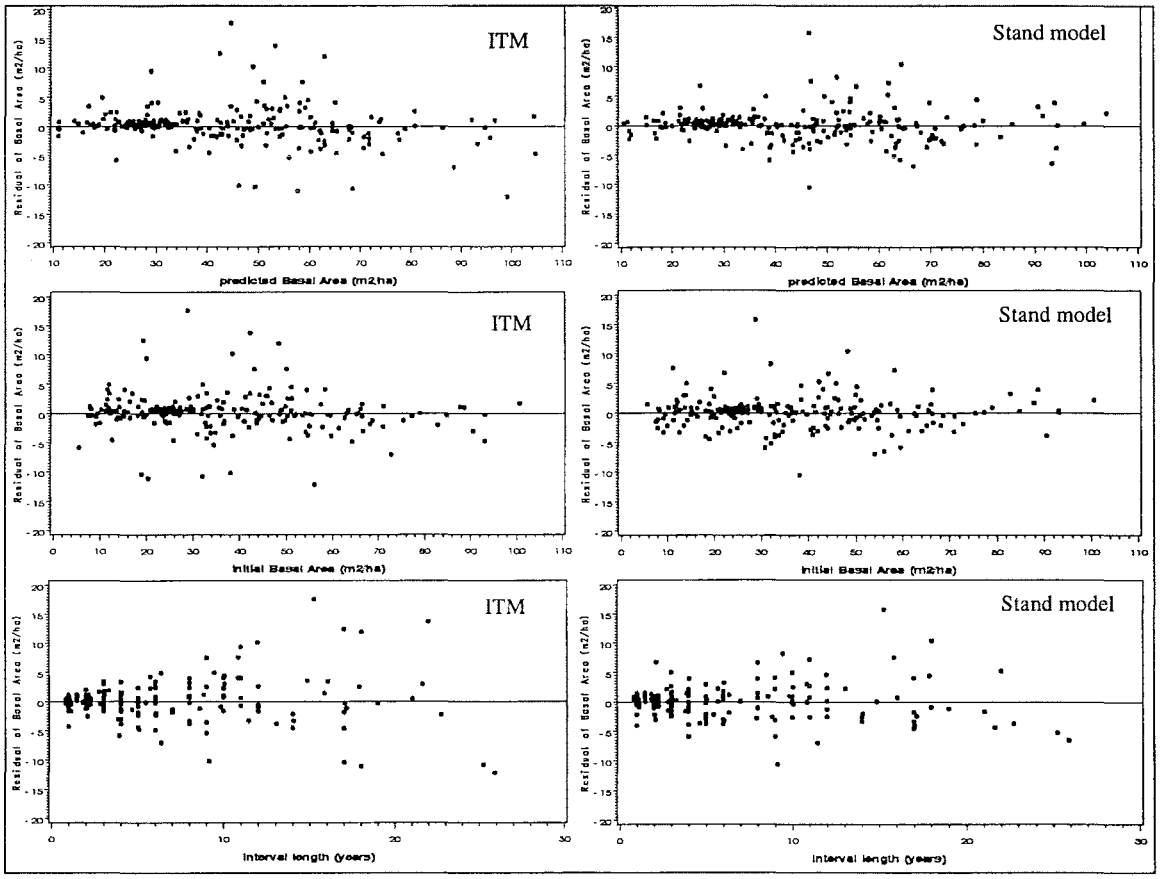


Figure 12.2 Basal area residuals (m^2/ha) against predicted values (top), initial basal area (middle) and projection interval (bottom), for the individual tree model (left) and stand level model (right) - *P. menziesii.*

For the New Zealand datasets, tree-level based estimations of basal area exhibited little bias (Table 12.2). However, basal area residuals from stand-level models were more tightly distributed around the zero reference line than residuals from ITMs (Figures 12.2 and 12.3). For the individual tree models, the magnitude of residuals increased with projection interval. Basal area projections up to 3 or 4 years ahead were virtually equal with both modelling approaches.

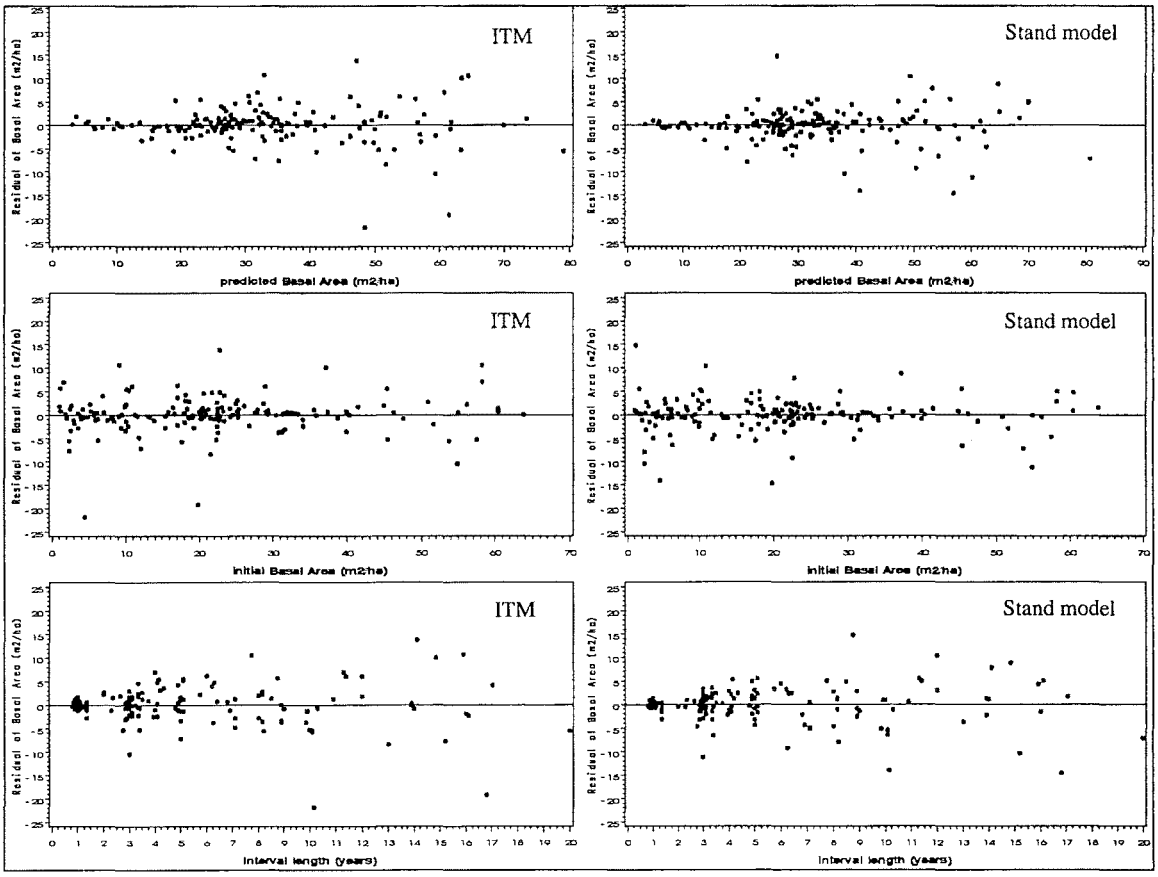


Figure 12.3 Basal area residuals (m^2/ha) against predicted values (top), initial basal area (middle) and projection interval (bottom), for the individual tree model (left) and stand level model (right) - *P. radiata.*

For all species, stand-level models provided stocking projections that were more accurate (lower mean residual) and precise (lower standard deviation of residuals) than those obtained from individual tree models (Table 12.3). Plots of residuals obtained from the two modelling approaches, against predicted stocking, initial stocking and interval length are displayed in Figures 12.4 to 12.6.

Table 12.3 Statistics of stocking residuals from unadjusted individual tree models (ITM) and stand models

Residual statistic	*E. grandis*		*P. menziesii*		*P. radiata*	
	ITM	Stand	ITM	Stand	ITM	Stand
Mean	29.82	-2.35	-22.01	-0.55	7.55	-0.08
Std. dev.	48.59	28.79	99.20	29.31	29.64	26.18
Skewness	1.22	-0.83	-5.31	-1.33	1.08	-0.21
Min. (1%)	-49.8	-90.2	-445.2	-115.4	-89.3	-97.3
Max. (99%)	188.1	57.1	54.7	71.5	94.0	76.4

For *E. grandis*, stocking projections from the ITM were biased, overestimating the mortality of young, highly stocked stands over longer projections (Figure 12.4). In these situations, the expansion factors (per hectare frequencies) of small trees were over-reduced at each iteration and these errors compounded. In order for the tree mortality model for *E. grandis* to be usable autonomously to simulate tree frequencies or stand stocking, it would need to be re-fitted using constraints to avoid this behaviour. However, this model can still be used in conjunction with the stand-level stocking model just to allocate predicted stand mortality across individual trees.

The extremely high mean residual of the ITM for *P. menziesii* was caused largely by one extreme negative residual (Figure 12.5). This residual arose from PSP 10410, for which initial stocking of 2373 stems/ha at age 19.25 was projected up to age 45.15. For that plot, actual stocking at age 45.15 was 1001 stems/ha, whereas predicted stockings from the ITM and the stand-level model were 1923 and 1014 stems/ha, respectively. Clearly, the individual-tree survival model for *P. menziesii* underestimated the mortality of highly stocked stands over long projection intervals. However, this model yielded unbiased predictions for stands stocked with up to 1000 stems/ha (Figure 12.5). Residual stocking statistics for Douglas-fir plots with stockings lower than or equal to 1000 stems/ha were 2.09, 20.03, -0.83, -89.4 and 70.9 for the statistics listed in Table 12.3, respectively.

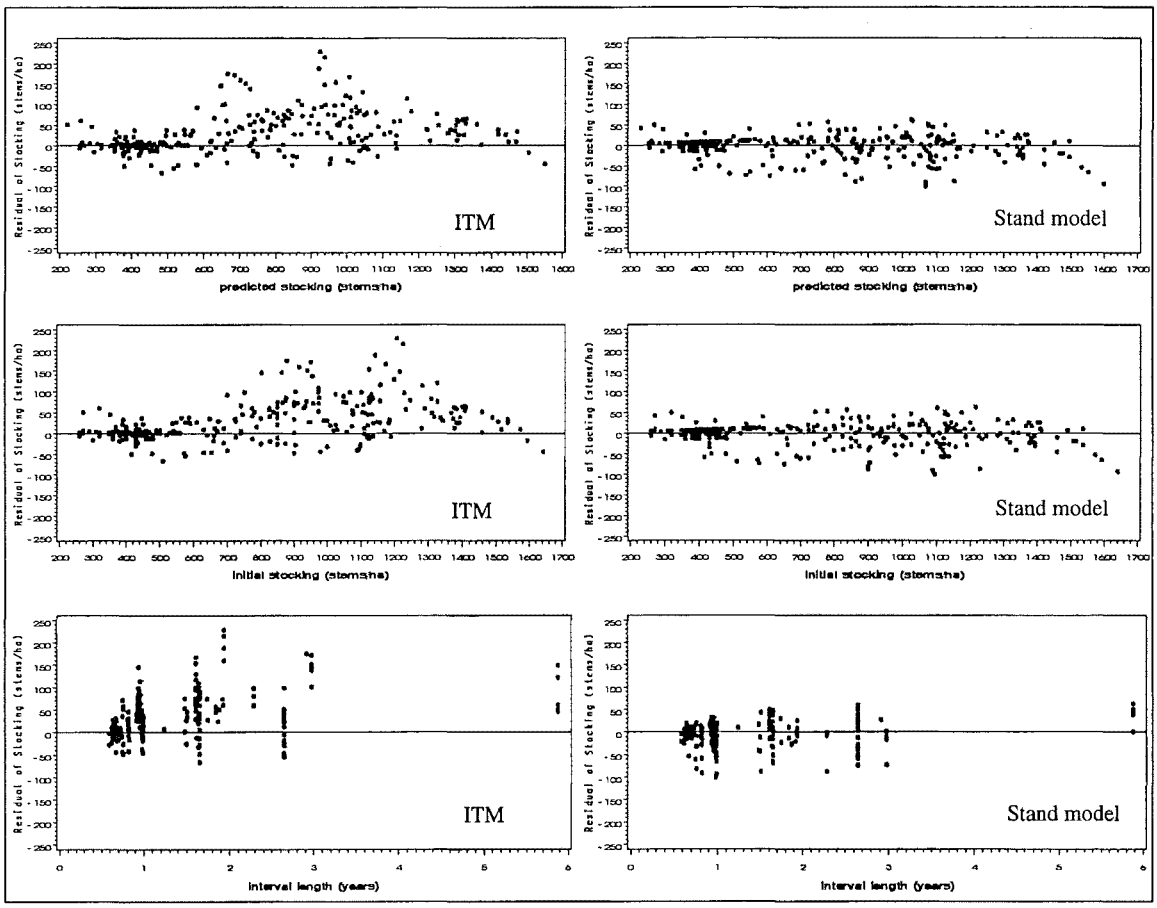


Figure 12.4 Stocking residuals (stems/ha) against predicted values (top) and initial stocking (bottom) for the individual tree model (left) and stand level model (right) - *E. grandis*.

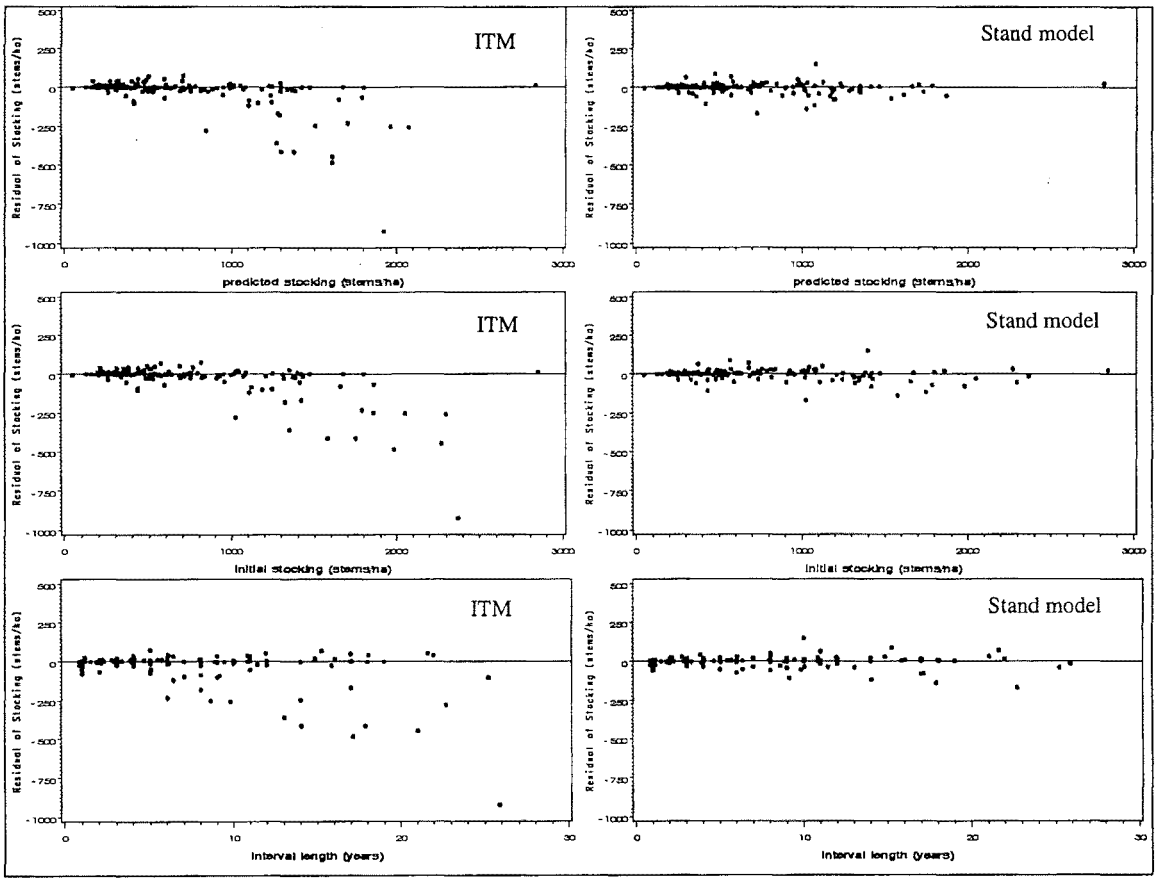


Figure 12.5 Stocking residuals (stems/ha) against predicted values (top), initial stocking (middle) and projection interval (bottom), for the individual tree model (left) and stand level model (right) - *P. menziesii.*

While the advantages of stand-level models over tree-level models for stocking predictions were clear for *E. grandis* and *P. menziesii*, they were less evident for *P. radiata* (Figure 12.6).

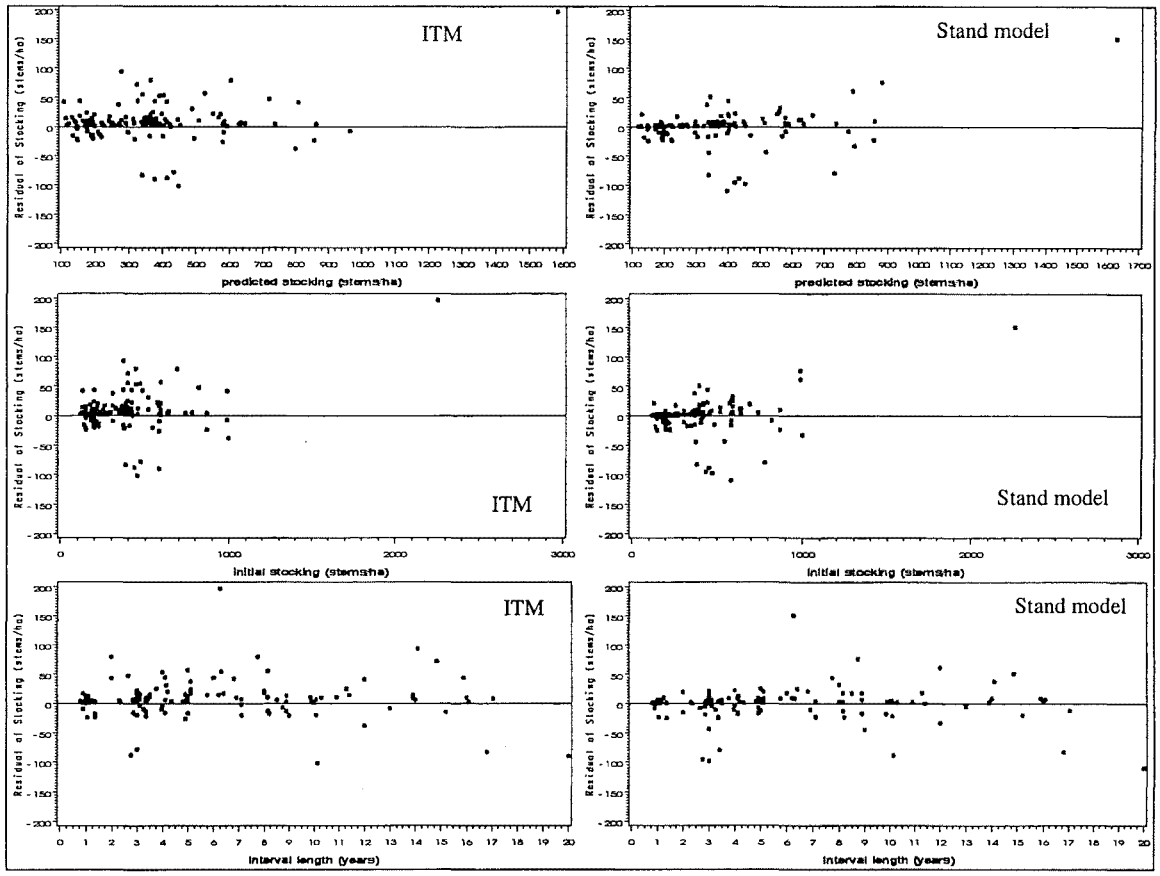


Figure 12.6 Stocking residuals (stems/ha) against predicted values (top), initial stocking (middle) and projection interval (bottom), for the individual tree model (left) and stand level model (right) - *P. radiata.*

12.3.2 Comparisons of predicted stand tables

Error index (equation 12.4) statistics for *E. grandis*, *P. menziesii* and *P. radiata* computed with different weighting factors are displayed in Tables 12.4, 12.5 and 12.6, respectively. For *E. grandis,* the unadjusted individual tree model (ITM) showed the lowest mean error index followed by the adjusted individual tree model (ITM_{adj}) regardless of the weighting factor (Table 12.4). The relative-basal-area-based disaggregative approach (RBA) and the diameter distribution model (DDM) ranked third or fourth depending on the weighting factor. Only the ITM and the DDM were found significantly different (according to the more exigent Tukey test) when some weighting factor was used (either tree basal area or tree volume).

Table 12.4 Means, standard deviations, minima and maxima of error indices computed with different weighting factors for *E. grandis*.

Weight	Significance	Method	Mean	LSD	Tukey	Std. dev.	Minimum	Maximum
tree volume	p=0.0068	ITM	88.4	b	b	55.8	12.6	493.7
		ITM_adj	91.9	b	ab	55.2	15.4	397.2
		RBA	102.4	a	a	61.4	16.1	417.8
		DDM	102.4	a	a	62.5	18.6	458.5
tree basal area	p=0.0041	ITM	8.93	c	b	4.3	1.9	37.7
		ITM_adj	9.27	bc	ab	4.2	2.0	32.2
		RBA	9.78	ab	ab	4.5	2.0	30.9
		DDM	10.15	a	a	3.9	2.4	26.8
none	p=0.0587	ITM	316.5	b	a	154.7	60.4	869.4
		ITM_adj	328.5	ab	a	158.6	61.7	854.2
		RBA	350.5	a	a	176.0	60.6	917.7
		DDM	343.2	a	a	144.2	76.3	783.0

Note: "Significance" obtained from the analysis of variance of error index for each weighting factor. Means with the same letter within each weighting factor are not statistically different according to the LSD or Tukey tests (α=0.05, error degrees of freedom=1084)

For *P. menziesii* when no weighting factor was used, the ITM_{adj} yielded the lowest mean error index, followed by the ITM and the RBA approach, while the DDM produced the largest mean error index (Table 12.5). When tree volume or tree basal area were included as weighting factors, the ranking of mean error indices was similar to the previous situation (no weighting factor), although the unadjusted ITM and the ITM_{adj} interchanged positions. However, the F tests did not detect a statistically significant effect of the modelling approach on the mean error index.

Table 12.5 Means, standard deviations, minima and maxima of error indices computed with different weighting factors for *P. menziesii*.

Weight	Significance	Method	Mean	Std. dev.	Minimum	Maximum
tree volume	p=0.2058	ITM	394.5	348.6	0.0	1894.1
		ITM_adj	408.1	359.8	12.8	1859.0
		RBA	418.0	360.8	12.7	1790.1
		DDM	463.4	363.4	21.2	2336.6
tree basal area	p=0.2005	ITM	19.37	14.0	0.0	78.9
		ITM_adj	19.68	13.5	1.5	68.1
		RBA	20.64	14.9	1.0	82.8
		DDM	21.92	11.9	3.0	65.1
none	p=0.1414	ITM	213.6	173.4	0.0	1357.0
		ITM_adj	208.3	148.1	25.9	1056.5
		RBA	230.4	185.7	25.1	1067.2
		DDM	239.9	130.8	34.9	1015.4

For the validation plots of *P. radiata*, the relative position of the four modelling approaches was constant regardless of the weighting factor, with the ITM_{adj} showing the lowest value, followed by the ITM, the RBA approach and finally the DDM. According to the F tests, there were no statistically significant effects of the modelling approach on mean error indices for any weighting factor (Table 12.6).

Table 12.6 Means, standard deviations, minima and maxima of error indices computed with different weighting factors for *P. radiata*.

Weight	Significance	Method	Mean	Std. dev.	Minimum	Maximum
tree volume	p=0.6416	ITM	225.8	220.0	0.3	1074.9
		ITM_adj	220.4	217.5	0.3	1057.7
		RBA	233.4	225.1	1.2	1116.2
		DDM	249.5	202.2	5.9	1037.1
tree basal area	p=0.3668	ITM	21.23	16.1	0.0	81.4
		ITM_adj	20.63	15.7	0.0	80.7
		RBA	21.76	16.5	0.1	79.6
		DDM	23.55	14.4	1.5	69.2
none	p=0.2318	ITM	183.6	98.8	0.3	696.6
		ITM_adj	180.1	111.5	0.4	1004.6
		RBA	196.9	159.1	1.2	1738.8
		DDM	203.5	90.9	61.2	862.2

For the New Zealand grown species there was a broad range of projection intervals, from one year up to 20 (*P. radiata*) or 26 years (*P. menziesii*), although the level of replication

for the longest intervals was limited. This allowed an analysis of the relative merit of the different modelling approaches depending on the length of the projection interval. In order to avoid confounding effects from tree-height and tree-volume models, error indices were weighted with tree basal areas. As expected, error indices increased with increasing projection interval (Figures 12.7 and 12.8). Decreases in model accuracy with increasing length of the projection period have been reported elsewhere (Shortt and Burkhart 1996; Rauscher *et al.* 2000).

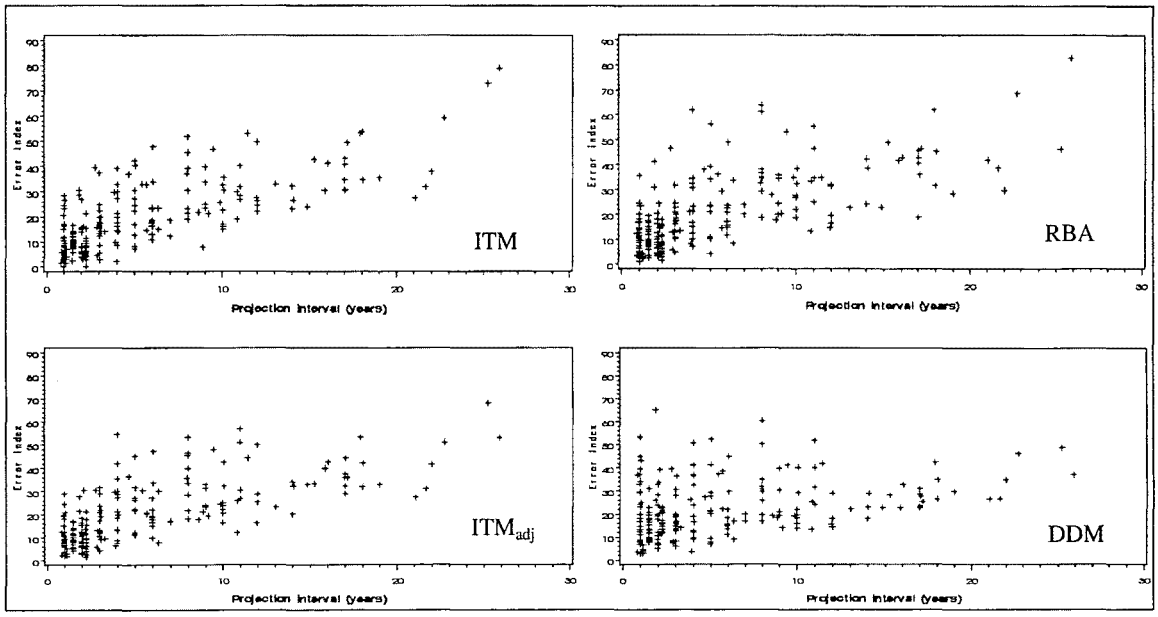


Figure 12.7 Basal-area-weighted mean error indices against projection interval (years) for all modelling approaches (*P. menziesii*).

The trend of increasing mean error indices with increasing projection interval was more pronounced with the unadjusted ITM (Figures 12.7 and 12.8). Conversely, this tendency was minimised with the DDM, suggesting that stand-level-based projections were more robust over long projection intervals. Those modelling approaches combining stand- and tree-level model components (i.e. ITM_{adj} and RBA) exhibited an intermediate behaviour. Comparing only unadjusted ITM and DDM (upper left and lower right graphs, respectively, in Figures 12.7 and 12.8) it is evident that ITMs generated lower error indices than DDMs for very short intervals (up to 3 to 4 years). For projections between 4 and 10 to 12 years both methods exhibited similar error indices, while DDMs showed lower error indices than ITMs over longer projections (greater than 12 years). These results are consistent with the findings reported by Shortt and Burkhart (1996) who

concluded that at shorter projection intervals the individual tree model performed better than stand-level models.

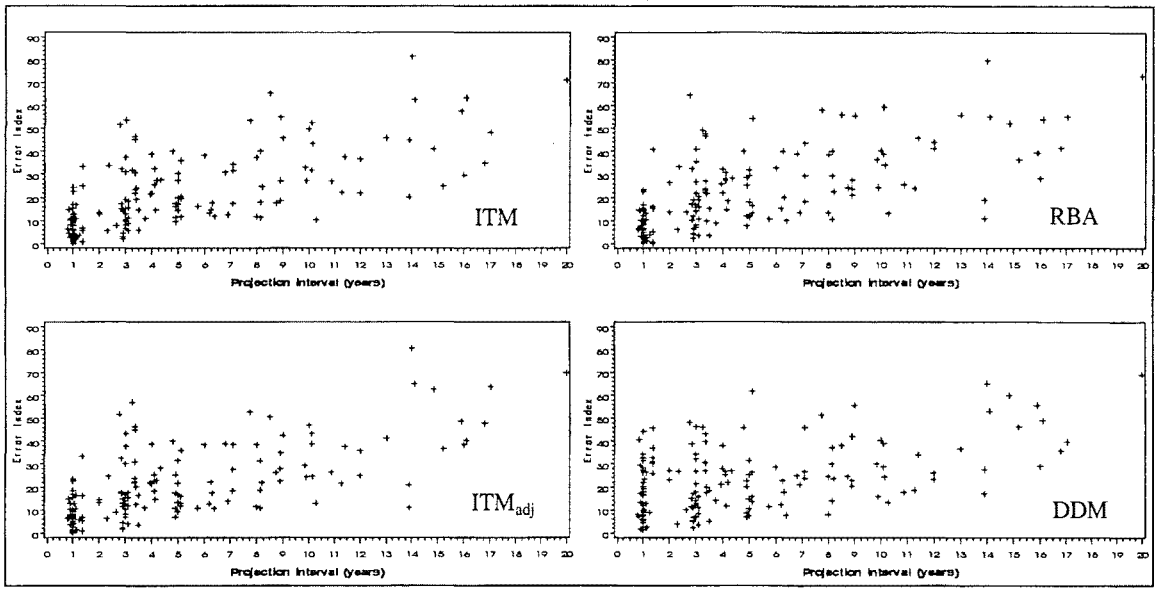


Figure 12.8 Basal-area-weighted error indices against projection interval (years) for all modelling approaches (*P. radiata*).

Finally, the predictive ability of the different modelling approaches was illustrated by plotting actual and predicted frequencies in three PSPs per species (Figures 12.9 to 12.11) as an example. For each species, the three PSPs included short (up to 6 years) and medium (up to 12 years) projection intervals as well as different stocking situations. In order to facilitate comparisons, the RBA approach was not included in Figures 12.9 to 12.11 because as Tables 12.4 to 12.6 show, this approach exhibited larger error indices than the other modelling approach combining stand-level and tree-level components (i.e. ITM_{adj}).

The graphs in Figures 12.9 to 12.11 indicate that, while all modelling approaches performed satisfactorily, the DDM produced stand tables in which the range of *dbh* classes always exceeded actual minima and maxima. In addition, the DDM was unable to depict irregular patterns in the stand tables, for example when there was an intermediate diameter class with no trees (e.g. *dbh* class 34 in PSP 8773, Figure 12.10). As pointed out in Section 9.8.3, the performance of this approach is very dependent on how accurately the maximum diameter is predicted. This is illustrated in Figure 12.11, where predicted reverse Weibull curves for PSPs 9285 and 9312 are left off-centred (i.e. maximum

diameter underpredicted). Conversely, the reverse Weibull curve for PSP 9501 in Figure 12.11 is right off-centred (i.e. maximum diameter overpredicted).

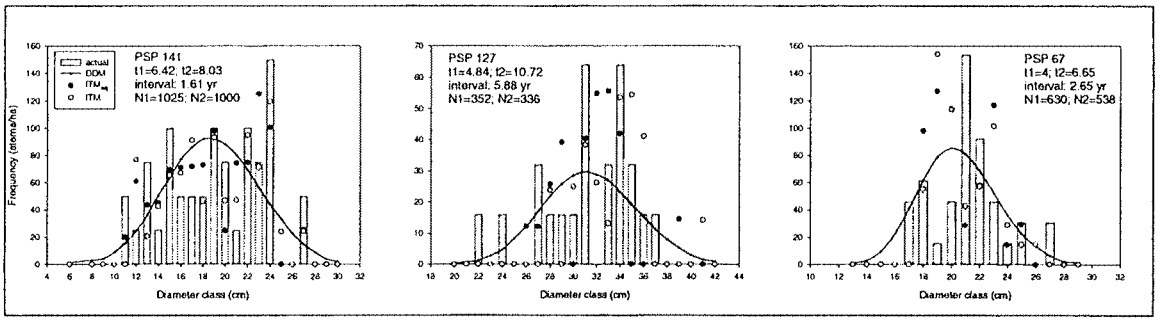


Figure 12.9 Actual and predicted frequencies (stems/ha) by diameter class (cm) from three modelling approaches (*E. grandis*).

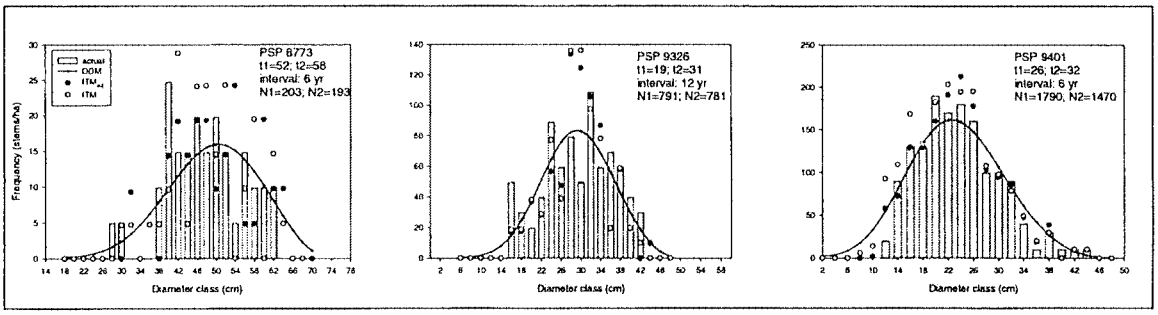


Figure 12.10 Actual and predicted frequencies (stems/ha) by diameter class (cm) from three modelling approaches (*P. menziesii*).

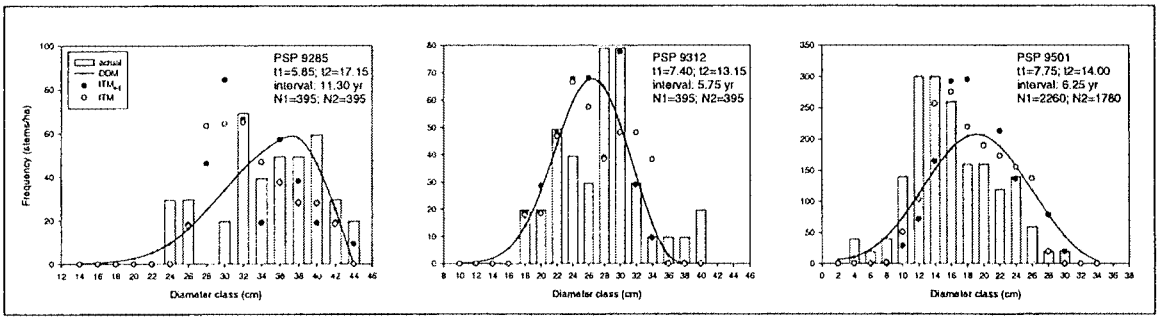


Figure 12.11 Actual and predicted frequencies (stems/ha) by diameter class (cm) from three modelling approaches (*P. radiata*).

12.4 DISCUSSION

Overall, the three modelling approaches that include tree-level components performed better than the DDM. These results are consistent with other studies (Borders and Patterson 1990; Knowe 1994; Knowe and Stein 1995) that obtained better representations of diameter distributions with the RBA approach (the latter two studies used relative diameter instead of RBA) than with DDMs. However, DDMs provided more accurate depictions of diameter distributions over long (greater than 12 years) projection intervals than other methods.

Modelling approaches including tree-level components can be separated into two groups. One group would be formed by the unadjusted ITM, which can be used as a stand-alone model. The other group is formed by the ITM_{adj} and the RBA dis-aggregative approach, both of which require stand-level models of basal area and stocking. A stand-level equation of MTH is required for all modelling approaches selected in this study[2] to estimate individual tree heights, which in turn are necessary to compute tree volumes.

Modelling approaches linked with stand-level model components have the advantage of being compatible with whole-stand and diameter distribution models. Within the two approaches of this group, results showed in this chapter indicated that the RBA approach was consistently outperformed by the ITM_{adj}, possibly because the diameter increment model relies on information from various tree and stand variables whereas the RBA approach does not. Knowe *et al.* (1997) also found that the individual tree model approach provided better representation of observed diameter distributions than the RBA dis-aggregative approach.

The unadjusted (ITM) of *P. radiata* was consistently outperformed by the adjusted individual tree model (ITM_{adj}), although the differences in error indices were not significant (Table 12.6). For *P. menziesii*, the ITM had lower error indices when tree volume or tree basal area were used as weighting factors (Table 12.5), whereas the ITM_{adj} exhibited a lower error index when no weighting factor was used (none of these differences was significant). Taking into account that stand-level-based projections of

[2] Pure ITM's without any reliance on a MTH equation are also feasible, although not explored in this study.

basal area and stocking for *P. menziesii* were more precise and accurate than those obtained from the unadjusted ITM (Tables 12.2 and 12.3), the ITM_{adj} approach was deemed the most appropriate for this species too.

For *E. grandis*, the unadjusted ITM provided the lowest error indices regardless of the weighting factor considered (Table 12.4). The ITM_{adj} ranked second and was not statistically different to the ITM for any case. Considering (i) the better performance of stand-level-based projections of stand basal area and stocking, as compared to tree-level based projections, and (ii) the compatibility advantage of the ITM_{adj}, this latter approach may be preferred for *E. grandis* as well.

12.5 CONCLUSIONS

For all three datasets, stand basal area and stocking projections from stand-level models exhibited lower means and standard deviations of residuals than projections from unadjusted individual tree models. Nonetheless, unadjusted individual tree models provided accurate estimations of stand basal area and stocking over short (up to 4 years) projection intervals. Stocking projections over long projection intervals from unadjusted individual tree models were extremely variable, which suggests that errors from tree-survival models readily compound.

The diameter distribution model exhibited the largest error indices for representing observed diameter distributions. This approach always predicted positive frequencies outside the range of observed diameter classes and was unable to depict irregular distribution patterns. Nevertheless, this approach provided accurate depictions of actual diameter distributions for most plots in Figures 9.28 to 9.30 and 12.9 to 12.11 and may be appropriate when a tree list is not available but information on maximum diameter and standard deviation (or variance) of diameters is. Another advantage of this approach (besides requiring less detailed input data than individual tree models) was its robustness over long projection intervals, for which lower error indices than other methods were exhibited.

Between those modelling approaches providing compatible tree-level and stand-level output resolution (i.e. RBA and ITM_{adj}) the adjusted individual tree model consistently showed lower error indices than the relative-basal-area-based dis-aggregative approach. This was attributed to the increased complexity of the ITM_{adj}, which relies on a number of site, stand and tree level variables, as opposed to the simple formulation of RBA models.

Adjusted individual tree models were preferred to unadjusted individual tree models for all species. For the New Zealand grown species, this decision was based on the better performance (lower error indices) of ITM_{adj} for depicting diameter distributions. For *E. grandis*, though, the main criterion for choosing the ITM_{adj} over the ITM was the compatibility with stand-level models, as the ITM exhibited slightly lower error indices than the ITM_{adj}.

PART III

GENERAL DISCUSSION AND SUMMARY OF CONCLUSIONS

CHAPTER 13

GENERAL DISCUSSION

Specific discussions about a broad range of topics ranging from taper modelling to detailed individual tree models have already been presented at the end of each of the preceding chapters. In this chapter, the new features of the study, an overview of the key findings in a general sense, and the areas where further research was considered relevant are discussed. Finally, the applicability, limitations and possible refinements of the proposed models are pointed out.

13.1 NEW FEATURES OF THE STUDY

13.1.1 Extensive comparison of existing taper equations and proposition of new variants

A large number of taper equations of various types were extensively evaluated and compared in a detailed fashion with two large datasets. The performances of these models were assessed in terms of diameter estimations as well as height and volume estimations. Such a thorough comparison of taper models has seldom been reported.

Variants of existing taper equations were proposed and they performed better than the original models. Furthermore, some of the proposed variants allowed the models to be

used for both under-bark and over-bark diameters, which was not feasible in the original models (Equations 1.3, 1.4, 1.8, 1.9 and 1.10). The modification of the classic segmented taper equation of Max and Burkhart (1976) provided more accurate and precise diameter, height and volume estimations than the original model. This model has the advantage that it can be explicitly integrated yielding a tree volume equation.

13.1.2 Evaluation of the composite approach

The approach of using a bark model to estimate under-bark diameters from predicted over-bark diameters (obtained from a taper equation fitted to over-bark data) has been used (Gordon *et al.* 1995, 1999) although not appropriately evaluated. The large dataset of measurements from *P. radiata* allowed this topic to be addressed. This analysis also provided useful insights into the sampling strategies to be used for collecting taper data.

13.1.3 Application of mixed effects analysis to checking parameter significance of taper equations

Although the autocorrelated nature of taper datasets is usually acknowledged, most studies simply accept the parameters obtained by least squares regressions. In this study, the significance of all parameters in the proposed taper and bark models was checked using mixed effects analysis. This technique has recently been used in a taper study by Tassisa and Burkhart (1998). An alternative method for checking the significance of parameters, consisting of extracting a single observation per tree, was used when mixed effects analysis was not feasible (Equation 2.11a).

Some parameters that were significant with ordinary least-squares techniques were not significant when autocorrelation was accounted for. Variables with non-significant parameters were removed or replaced, resulting in more stable and robust models.

13.1.4 Extensive comparisons of growth modelling approaches

The availability of three datasets, two of which were very large, permitted extensive comparisons of differing approaches to growth and yield modelling. These comparisons were performed at the following levels:

(i) Whole stand models. Numerous equation forms were compared for each variable. A polymorphic form derived from the Chapman-Richards model that has seldom been used in other studies (an exception being Amaro *et al.* 1998) exhibited the best fit for mean top height in the three species.

(ii) Diameter distribution models. The use of skewness information with the reverse Weibull distribution was explored, and severe inconveniences that militated against its use were described (see Section 9.2). Equations for estimating the standard deviation of diameters after a thinning operation were developed. This model component, largely unreported in the literature, is considered crucial for the implementation of diameter distribution models.

(iii) Individual tree models. Diameter increment models were compared with diameter difference equations and a relative-basal-area-based approach. Datasets with different structures were formed for these comparisons.

(iv) Models for individual tree heights. Extensive comparisons of basic equation forms and modelling approaches were performed for this component.

13.1.5 Development of a growth and yield model for *E. grandis* in Uruguay

The taper and growth and yield models built in this study for *E. grandis* growing in Uruguay constitute a new development as there are no models for that species (or for any other) publicly available in the country. The growth and yield model developed here will allow more informed decisions to be made, which in turn may affect current silvicultural

and managerial practices. It should be noted, though, that the model is provisional and should be revised when more PSP data become available.

13.1.6 Development of an individual tree model for Douglas-fir in New Zealand

Although there are stand-level models for Douglas-fir in New Zealand, to the best of the author's knowledge, the individual tree model developed in this study will be the first. The proposed individual tree model will enable more accurate representations of future stands, particularly over short (up to 5 year) projections.

13.2 OVERVIEW OF THE KEY FINDINGS

13.2.1 Taper modelling

In terms of diameter, height and volume estimation capabilities, variable exponent models exhibited the maximum accuracy and precision amongst all models tested. This was consistent with results reported by Newnham (1992) and Kozak and Smith (1993). However, these models cannot be inverted to estimate heights, nor can they be explicitly integrated to calculate volumes. A modified version of the segmented taper equation (Max and Burkhart 1976) was proposed. This model lacks the disadvantages of variable-exponent models, yet provides comparable levels of accuracy and precision.

As in previous taper studies (Newnham 1992; Kozak and Smith 1993; Muhairwe 1999) it was found that the selection of taper estimation systems should be based on a detailed sequence of analyses that consider the accuracy and precision of diameter, height and tree-volume estimations. These analyses should be performed by relative-height and diameter classes as well as at the overall level. Only with this approach would it be possible to choose the best attainable taper estimation systems for operational applications.

Fitting a bark model from about 150 trees or 2500 measurements and using this equation to estimate under-bark diameters from predicted over-bark diameters produced virtually the same residual statistics as estimating under-bark diameters directly. Therefore, when collecting data for taper modelling, only a proportion of trees would need to be measured for under-bark diameters, thus saving resources. Given that the lower part of the stem exhibited the largest variation in terms of bark thickness (and it is the most valuable) it is suggested to measure under-bark diameters (or bark thickness) at breast height and at half breast height (0.65 or 0.70 m) before felling each tree.

Given the considerable level of autocorrelation in taper datasets, the use of mixed effects analysis was a powerful technique to check the significance of parameters. This was particularly relevant for variable exponent models, for which this technique allowed the most meaningful set of variables to be selected.

13.2.2 Modelling stand structure and dynamics

Some equation forms consistently exhibited the lowest MSE and mean residual for certain variables across datasets. A polymorphic Chapman-Richards derived by letting the *shape* parameter be the free parameter (the most common polymorphic version is derived from the *rate* parameter, as explained in Section 7.4.3.2) showed the best fit for modelling mean top height (MTH) development over time. This variant of the Chapman-Richards model was also preferred by Amaro *et al.* (1998) for modelling MTH growth of eucalypt species in Portugal.

The well-known Schumacher polymorphic equation was consistently selected as the base model for depicting stand basal area growth and yield. This is consistent with the findings by Woollons and Wood (1992) who found that the Schumacher and Weibull equations outperformed other equations for basal area projections across various datasets. Exponential regressions including stand basal area, mean top height, age and stocking, fitted by weighted non-linear regression, exhibited maximum levels of accuracy and precision for stand volume estimations. Amateis *et al.* (1986) selected a linear regression for estimating the logarithm of stand volume from basal area, stocking, MTH and age, which is an equation form equivalent to the exponential regressions adopted in this study.

The effect of altitude was successfully accounted for in basal area and MTH equations for the New Zealand grown species. The *E. grandis* model included dummy variables to account for locality effects, a procedure successfully used in other modelling studies (Temu 1992; Whyte *et al.* 1992; Lee 1998). Thinning effects on basal area, interpreted as the difference in growth between two stands with identical basal area and age, one just thinned and one unthinned, varied with species. Basal area growth of *P. radiata* was reduced by thinning, which was attributed to the loss of crown surface. For *P. menziesii*, basal area after thinning was positively affected by thinning, which may be caused by the removal of sub-dominant trees that are likely to be more affected by *Phaeocryptopus* than dominant trees. For both species, the effect of thinning was more pronounced with increasing proportions of basal area thinned and less pronounced as thinning age increased. A slightly negative effect of thinning on basal area growth was detected for *E. grandis*. However, limitations in the amount of data from thinned plots precluded incorporating a thinning index in the basal area equation for this species.

Diameter distributions obtained with the reverse Weibull parameter-recovery approach depicted diameter distributions in reasonable agreement with actual distributions in most example plots (Figures 9.28-9.30 and 12.9-12.11). For the New Zealand grown species, this method generated lower error indices over long (greater than 12 years) projection intervals than all other methods (Figures 12.7 and 12.8). This confirmed the merits of this approach that were pointed out by Xu (1990); Kuru *et al.* (1992) and Whyte and Woollons (1992). The location parameter of the reverse Weibull distribution was defined by increasing predicted maximum diameters by a fixed percentage that varied across datasets. This procedure provided diameter distribution depictions that were closer to actual stand tables than distributions obtained without adjusting predicted maximum diameters. Although this empirical solution to the problem of defining the population extreme (as compared to the sample extreme) provided good empirical results, a theoretical study along the lines of Xu's (1990) work would be required to clarify this topic.

Diameter increment equations provided more accurate and precise estimations of diameter growth than difference equations. Nonetheless, evaluation of the latter equations was constrained by the need to maintain their path-invariance property. This

invalidated the inclusion of numerous tree and stand variables that were crucial for diameter increment equations. Individual tree models based on equations describing the trend of relative basal area over time were not as accurate as models based on diameter increment equations, which is consistent with the study by Knowe *et al.* (1997). The main advantage of the relative-basal-area approach is that it ensures compatibility between tree-level and stand-level outputs. While this is not implicitly assured in the diameter-increment approach, it can be forced by adjusting predicted diameters and frequencies at each iteration so that aggregated values of basal area and stocking equal stand-level estimates.

In terms of diameter distribution depiction, the adjusted individual tree model approach provided the lowest error indices for the New Zealand datasets. The unadjusted individual tree model exhibited slightly lower error indices for *E. grandis*. Nonetheless, the adjusted individual tree model was also recommended for *E. grandis* in order to avoid incompatibility with stand-level models, which may be misleading for the user. These findings suggest that the better performance of stand-alone individual tree models over stand-level-driven dis-aggregative approaches reported in previous studies (Knowe *et al.* 1997; Ritchie and Hann 1997a) could be attributed mostly to the features of the datasets used (small datasets, short intervals) rather than the modelling approach.

This study has demonstrated that, in order to obtain maximum levels of accuracy and precision for predicting individual tree heights, a detailed and multi-stage process should be followed. Several equation forms depicting the relationship between tree diameters and heights were compared. Once a few equations were selected, the parameter-prediction and asymptote-prediction approaches (see Section 11.2) were tried and the models evaluated by diameter class. The Petterson equation, which is the standard height/diameter model in New Zealand, was selected for *P. menziesii* but not for *P. radiata*. The asymptote prediction approach, which was selected for *P. radiata*, was simpler than the parameter prediction approach but it can only be used when the relationship between mean top height and mean top diameter is clearly defined.

The proposed models provide a range of options for performing forecasts of future stand conditions depending on the type of input data available (Table 13.1). All these options (i.e. whole-stand, diameter distribution and individual tree models) are fully compatible.

The models were programmed in Visual Basic for Applications (VBA) in Microsoft Excel spreadsheets, a familiar user-friendly environment (Appendix 4). Provisions were made for merchantable volume estimations by log type using a taper equation developed in Chapter 3 for *E. grandis* and existing taper equations for *P. radiata* and *P. menziesii*. Whole-stand models were also programmed in Javascript by Dr. Euan Mason and have been available at the New Zealand School of Forestry web site as of January 2001 (www.fore.canterbury.ac.nz, under the "Software" tab).

Table 13.1 Model options according to the level of detail in the input data

Input data	Model type	Output resolution (at thinnings and harvest)
Stand statistics (basal area, MTH, stocking, age[1])	Whole stand model	Basal area, MTH stocking and stand volume
Stand statistics plus maximum plot diameter and standard deviation (or variance) of diameters	Diameter distribution model	As above plus number of trees by diameter class and merchantable volumes by log type[2]
Tree list: tree diameters (required) and tree heights (optional)	Individual tree model	Stand statistics plus tree list and merchantable volumes by log type

[1] Information on altitude (New Zealand) or zone (Uruguay) also required for all models.

[2] Merchantable volumes by log type obtained with taper equations.

13.3 AREAS FOR FURTHER RESEARCH

Development of the models described in this thesis revealed opportunities for further research on a few topics. The following topics were identified:

(i) development of taper equations with the flexibility and accuracy of variable exponent models that can also be inverted for height estimations and explicitly integrated for volume estimations;

(ii) development of a method for defining the maximum stand diameter (location parameter of the reverse Weibull distribution). This method needs to be theoretically sound but it must also be possible to implement it in a growth and yield simulator (see Sections 7.7.1 and 9.7 and 9.10); and

(iii) evaluation of diameter models based on difference equations augmented with the same type of tree and stand variables that were used for diameter increment models. This would imply losing the path-invariance of these equations but increasing substantially their predictive ability. According to the results reported in Section 10.4.3, anamorphic equations would be more appropriate than polymorphic equations.

(iv) comparisons of the models reported here against existing models with comparable levels of resolution for the CNI region. At the stand-level, PPM88 (*P. radiata*) and DFCNIGM or DFEARLY (*P. menziesii*) may be compared against the new whole-stand and diameter distribution models. At the tree-level, the model developed by Gordon and Shula (1999) may be compared with the tree-level model reported here for *P. radiata*.

13.4 APPLICABILITY, LIMITATIONS AND POSSIBLE REFINEMENTS OF THE PROPOSED MODELS

The models for *E. grandis* were constructed with data from Zones 7, 8 and 9 of Uruguay and therefore, their use should be limited to these zones. The database available for this species was very limited, particularly in terms of age coverage (the oldest plots were aged 16 years) and thinnings. The number of thinned plots was limited and most thinned plots received only one thinning operation. Data from Zones 8 and 9 were scarcer than data from Zone 7. In view of all these limitations, the models are considered provisional and should be revised when more PSP data become available.

The New Zealand datasets were very complete and had excellent coverage for most relevant attributes (age, altitude, site index, thinning regimes) particularly for *P. radiata*. The plots were located in the Central North Island of New Zealand and the majority were from Kaingaroa forest. The models are obviously recommended for this region unless they can be validated for similar regions and show acceptable performance. The equations for the main stand-level components (MTH and basal area) included altitude,

which provided enhanced sensitivity to these models as compared to existing models (PPM88 and DFCNIGM3). This variable is readily available from topographic regional maps.

If tree lists are available, adjusted individual tree models are recommended to achieve maximum levels of flexibility and accuracy for forecasting stand structure and dynamics over short to medium projection intervals. However, for long-interval projections (i.e. more than 12 years for the New Zealand grown species and possible more than 6 to 8 years for *E. grandis* in Uruguay), diameter distribution models may provide more robust and accurate projections.

The effect of pruning on the basal area growth of *P. radiata* was evaluated and found negligible. However, pruning information was limited and most pruning operations in the dataset were performed at the same time as thinning operations. The negative effect of thinning on basal area growth found for this species may subsume any detrimental effect of pruning on basal area growth. Greater sensitivity to extreme thinning and pruning regimes may be achieved with EARLY, a model specifically conceived to handle these effects. EARLY outputs obtained after tending operations are performed can be used as starting values for the stand-level model developed in this study.

No information on pruning was available for *P. menziesii*, as this species is typically grown without pruning operations in New Zealand. However, if intensive thinning or pruning operations are to be simulated, the DFEARLY model may be used for the initial phase of the rotation, where tending activities are normally concentrated.

When developing basal area and mean top height equations for *P. radiata*, residuals were plotted against GF rating to ascertain whether improved breeds exhibited different growth patterns. No trends were visible, although the range of GF ratings available in the dataset was limited and unbalanced. It should be noted that the effect of GF rating was evaluated with difference equations where the initial yield determines to a large extent the predicted growth ahead. Therefore, having found no effects of GF rating on basal area and stand height growth does not mean that stands planted with improved breeds would not accrue larger yields.

CHAPTER 14

SUMMARY OF CONCLUSIONS

The main conclusions for the different topics covered in this study were already highlighted and discussed at the end of each chapter. The key findings were also discussed in Chapter 13. In this chapter, the main conclusions for each topic are briefly enumerated. Finally, all selected equations and corresponding parameter estimates are listed in Tables 14.1 to 14.4.

14.1 TAPER AND BARK MODELS

- Variable exponent models exhibited maximum levels of accuracy and precision for diameter, height and volume estimations. Mean over-bark diameter, merchantable height and tree-volume residuals were, respectively, -0.04 cm, -0.055 m and -0.002 m^3 for *E. grandis* (Equation 2.14a) and -0.043 cm, 0.085 m and 0.001 m^3 for *P. radiata* (Equation 1.16).

- A variant of the Max and Burkhart (1976) segmented taper equation that can be inverted for height estimations and explicitly integrated for volume calculations was proposed and used for *E. grandis*. This equation provided reasonable levels of accuracy with mean over-bark residuals of -0.052 cm, -0.181 m and -0.002 m^3 for diameter, merchantable height and tree volume, respectively.

- Mixed effects analysis allowed selecting the set of variables that were really significant when autocorrelation was accounted for, resulting in more robust and stable taper and bark equations.

- Under-bark diameter estimations obtained by (i) predicting over-bark diameters with a taper equation fitted to over-bark data; and (ii) applying bark models to predicted diameters in (i), generated residual statistics that were almost equal to residual statistics obtained by fitting taper equations directly to under-bark data.

- Using a sub-sample of about 170 trees (40% of trees available) with under-bark measurements generated almost the same residual statistics of under-bark diameter estimations through the composite approach than using all (428) trees.

Table 14.1 Listing of selected taper and bark models

	E. grandis		*P. radiata*	
Model component	Equation	Parameters	Equation	Parameters
Over-bark taper equation	2.14a* 2.11a	Table 5.1 Table 5.2	1.16	Table 5.1
Under-bark taper equation	2.11a	Table 5.2	1.16	Table 5.1
Bark models	2.25a	Table 4.5	2.25b	Table 4.5

14.3 MODELLING STAND STRUCTURE AND DYNAMICS

- Equation 9.1, a polymorphic Chapman-Richards equation that has not been widely used, exhibited the best fit (as judged mainly by the MSE, mean residual and graphical plots of residuals), for mean top height projections for the three species.

- Equation 9.2, the well-known Schumacher polymorphic equation, showed the best fit for basal area predictions with the three datasets.

- Exponential regressions (Equation 9.11) including inputs of stand basal area, mean top height, age and stocking, fitted by weighted non-linear regression, provided maximum levels of accuracy and precision for stand volume estimations.

- Stand basal area and stocking were more accurately and precisely estimated with stand-level models than with stand-alone individual tree models. In the latter models, errors of individual trees easily compounded, particularly for stocking estimations.

- Individual tree models based on diameter increment equations and logistic regressions for estimating probabilities of tree survival, provided the lowest error indices for depicting diameter distributions, compared to the relative-basal-area dis-aggregative approach and diameter distribution models.

- Unadjusted individual tree models produced slightly lower error indices than adjusted individual tree models for *E. grandis*. Adjusted individual tree models generated the lowest error indices for the New Zealand datasets, being more robust over long projections. Considering this and the compatibility aspect, adjusted individual tree models were recommended for all species.

- Error indices obtained with the reverse Weibull parameter-recovery approach were always the largest, although not significantly different to error indices obtained with other methods in most cases. This suggested that the reverse Weibull approach may also be a useful technique for generating diameter distributions when tree lists are not available for model development. This method exhibited lower error indices over long projection intervals than all other methods.

- Different equation forms describing the relationship between tree diameters and heights were selected for each species. Maximum levels of accuracy and precision for tree height estimations were achieved with the parameter prediction approach for *E. grandis* and *P. menziesii*, and with the asymptote prediction approach for *P. radiata*. The latter approach may be applicable only when the relationship between MTH and MTD is unambiguously defined.

Table 14.2 Listing of selected equations for the proposed growth model for *E. grandis*

Model component	Units	Equation	Parameters	Remarks
Mean top height (MTH)	m	9.1	Table 9.1	Chapman-Richards polym.
Stand basal area (G)	m^2/ha	9.2	Table 9.3	Schumacher polymorphic
Stocking (N)	stems/ha	9.8	Table 9.10	Clutter *et al*. (1983)
Stand volume (V)	m^3/ha	9.11	Table 9.12	Exponential regression
Maximum diameter (D_{max})	cm	9.3	Table 9.5	Chapman-Richards polym.
Standard deviation of dbh (D_{std})	cm	9.6	Table 9.7	Gompertz polymorphic
Parameters a, b and c (reverse Weibull)		Procedure detailed in Section 9.7		
Diameter increment (ΔD)	cm	in Table 10.11	Table 10.11	Linear regression
Probability of tree survival (P_{live})	-	10.3	Table 10.4	Logistic model
Individual tree heights (h)	m	11.16	Table 11.10	Parameter pred. approach

Table 14.3 Listing of selected equations for the proposed growth model for *P. menziesii*

Model component	Units	Equation	Parameters	Remarks
Mean top height (MTH)	m	9.1	Table 9.1	Chapman-Richards polym.
Stand basal area (G)	m^2/ha	9.2	Table 9.3	Schumacher polymorphic
Stocking (N)	stems/ha	9.9	Table 9.10	Clutter and Jones (1980)
Stand volume (V)	m^3/ha	9.11	Table 9.12	Exponential regression
Maximum diameter (D_{max})	cm	9.4	Table 9.5	Monomolecular polym.
Standard deviation of dbh (D_{std})	cm	9.7	Table 9.7	Chapman-Richards polym.
Parameters a, b and c (reverse Weibull)		Procedure detailed in Section 9.7		
Diameter increment (ΔD)	cm	10.4	Table 10.13	Exponential regression
Probability of tree survival (P_{live})	-	10.3	Table 10.5	Logistic model
Individual tree heights (h)	m	11.5	Table 11.11	Parameter pred. approach

Table 14.4 Listing of selected equations for the proposed growth model for *P. radiata*

Model component	Units	Equation	Parameters	Remarks
Mean top height (MTH)	m	9.1	Table 9.1	Chapman-Richards polym.
Stand basal area (G)	m^2/ha	9.2	Table 9.3	Schumacher polymorphic
Stocking (N)	stems/ha	9.10	Table 9.10	Clutter and Jones (1980)
Stand volume (V)	m^3/ha	9.11	Table 9.12	Exponential regression
Maximum diameter (D_{max})	cm	9.5	Table 9.5	Schumacher polymorphic
Standard deviation of dbh (D_{std})	cm	9.7	Table 9.7	Chapman-Richards polym.
Parameters a, b and c (reverse Weibull)		Procedure detailed in Section 9.7		
Diameter increment (ΔD)	cm	10.5	Table 10.13	Growth x logistic modifier
Probability of tree survival (P_{live})	-	10.3	Table 10.6	Logistic model
Individual tree heights (h)	m	11.12; 11.20	Table 11.12	Asymptote pred. approach

Acknowledgements

I would like to thank Dr. Euan Mason, my supervisor, for his encouragement, advice and support throughout all the stages of this study. Dr. Richard Woollons also provided valuable assistance as well as the tree taper data of radiata pine. Thanks are also due to Dr. Bruce Manley, who provided useful advice on various specific topics of the research.

The supply of the growth modelling data for the New Zealand grown species by Fletcher Challenge Forests is gratefully acknowledged. The data were extracted from *Forest Research*'s PSP system by Ms. Judy Hayes, whose help and positive attitude to clarify all my queries is thanked. Thanks are also due to the four companies that supplied PSP data of *E. grandis* in Uruguay.

Financial support from the New Zealand Official Development Assistance (NZODA) division of the Ministry of Foreign Affairs and Trade, which made possible this study, is deeply appreciated. I would like to thank Dr. John Pickering for his competent handling of all administrative matters regarding my NZODA Study Award.

Thanks are also due to Mr. Robert Shula (*Forest Research*) for his guidance at early stages of development of individual tree models, and to Mr. Gustavo Balmelli (National Forestry Programme, INIA Uruguay) for his suggestions on the design and improvement of the spreadsheet-based simulators (CITASS) and their extensive beta testing. The valuable amendments and suggestions provided by the external examiners, Dr. A.G.D Whyte and Prof. H.E. Burkhart are sincerely appreciated.

Special thanks are due to the rest of the staff at the School of Forestry and postgraduate fellows for their friendship and support. Specially, I would like to thank my colleague Mike Watt for his thorough revision of the literature review chapter and his useful suggestions.

I thank my parents for their unwavering faith and encouragement and for their selfless effort in my early education.

Finally, I thank Isabel, my wife and friend, and our beloved daughter Valentina, for their endless patience, love and encouragement.

To anyone I may have inadvertently forgotten, sorry, and thanks.

References

Alder, D. 1979. A distance-independent tree model for exotic conifer plantations in East Africa. Forest Science 25(1): 59-71.

Alemdag, I.S. 1988. A ratio method for calculating stem volume to variable merchantable limits, and associated taper equations. The Forestry Chronicle 64(1): 18-26.

Allen, P.J., N.B. Henry and P. Gordon. 1992. Polynomial taper model for Queensland plantation hoop pine. Australian Forestry 55: 9-14.

Amaro, A., D. Reed, M. Tomé and I. Themido. 1998. Modeling dominant height growth: *Eucalyptus* plantations in Portugal. Forest Science 44(1): 37-46.

Amateis, R.L., H.E. Burkhart and T.E. Burk. 1986. A ratio approach to predicting merchantable yields of unthinned loblolly pine plantations. Forest Science 32(2): 287-296.

Amateis, R.L., H.E. Burkhart and T.A. Walsh. 1989. Diameter increment and survival equations for loblolly pine trees growing in thinned and unthinned plantations on cutover site-prepared lands. Southern Journal of Applied Forestry 13(4): 170-174.

Amidon, E.L. 1984. A general taper functional form to predict bole volume for five mixed-conifer species in California. Forest Science 30(1): 166-171.

Arabatzis, A.A. and H.E. Burkhart. 1992. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations. Forest Science 38(1): 192-198.

Assmann, E. 1970. The principles of forest yield study. Permagon Press, Oxford. 506 pp.

Avila, O.B. and H.E. Burkhart. 1992. Modeling survival of loblolly pine trees in thinned and unthinned plantations. Canadian Journal of Forest Research 22: 1878-1882.

Bailey, R.L. 1994. A compatible volume-taper model based on the Schumacher and Hall generalized constant form factor volume equation. Forest Science 40(2): 303-313.

Bailey, R.L. and J.L. Clutter. 1974. Base-age invariant polymorphic site index curves. Forest Science 20: 155-159.

Bailey, R.L. and T.R. Dell. 1973. Quantifying diameter distributions with the Weibull function. Forest Science 19(2): 97-104.

Bailey, R.L. and K.D. Ware. 1983. Compatible basal-area growth and yield model for thinned and unthinned stands. Canadian Journal of Forest Research 13: 563-571.

Battaglia, M. and P.J. Sands. 1998. Process-based forest productivity models and their application in forest management. Forest Ecology and Management 102: 13-32.

Beekhuis, J. 1966. Prediction of yield and increment in *Pinus radiata* stands in New Zealand. New Zealand Forest Service. Technical Paper. No. 49.

Belton, M. 2000. Commercial Douglas fir forestry in New Zealand. New Zealand Tree Grower 21(3): 35-37.

Berkey, C.S. 1982. Bayesian approach for non-linear growth models. Biometrics 38: 953-961.

Biging, G.S. and M. Dobbertin. 1995. Evaluation of competition indices in individual tree growth models. Forest Science 41(2): 360-377.

Bitoki, O., L.R. Gering, T.B. Lynch and P.A. Murphy 1998. An individual tree basal area growth model for uneven-aged stands of shortleaf pine (*Pinus echinata* Mill.) in the Ouachita Mountains in Arkansas and Oklahoma. In: T. A. Waldrop (Ed.). Proceedings of the Nineth Bienn. South. Silv. Res. Conf. USDA Forest Service. Gen. Tech. Report SRS-20. p 522-527

Bliss, C.I. and K.A. Reinker. 1964. A log-normal approach to diameter distributions in even-aged stands. Forest Science 10: 350-360.

Borders, B.E., R.L. Bailey and M.L. Clutter. 1987. Forest growth models: parameter estimation using real growth series. In: Ek, S.R., Shifley, S.R. and Burk, T.E. (Editors). Forest Growth modelling and Prediction, Proceedings IUFRO Conference, Mineapolis, MN, USDA Forest Service. General Technical Report. NC-120. 660-667 pp.

Borders, B.E., R.L. Bailey and K.D. Ware. 1984. Slash pine site index from a polymorphic model by joining (splining) nonpolynomial segments with an algebraic difference method. Forest Science 30(2): 411-423.

Borders, B.E. and W.D. Patterson. 1990. Projecting stand tables: a comparison of the Weibull diameter distribution method, a percentile-based projection method and a basal area growth projection method. Forest Science 36(2): 413-424.

Bruce, D., R.O. Curtis and C. Vancoevering. 1968. Development of a system of taper and volume tables for red alder. Forest Science 14(3): 339-350.

Bruce, D. and L.C. Wensel 1987. Modelling forest growth: Approaches, definitions and problems. In: A. R. Ek, S. R. Shifley and T. E. Burk (Ed.). Proceedings of the IUFRO symposium on forest growth modelling and prediction. Minneapolis, Minesota. USDA Forest Service. General Technical Report NC-120. 1-8

Bunnell, F.L. 1989. Alchemy and uncertainty: What good are models? USDA Forest Service. General Technical Report. PNW-GTR-232. 27 pp.

Burgess, I.P. 1988. Provenance trials of *Eucalyptus grandis* and *E. saligna* in Australia. Silvae Genetica 37: 221-227.

Burk, T.E. and H.E. Burkhart. 1984. Diameter distribution and yield of natural stands of loblolly pine. School of Forestry and Wildlife Resources. Virginia Polytechnic ans State University. FWS-1-84. 46 pp.

Burkhart, H.E. 1987. Data collection and modelling approaches for forest growth and yield prediction. In: Predicting forest growth and yield - Current Issues, Future prospects. Institute of Forest Resources. University of Washington. Contribution Nr. 58. 3-16

Burkhart, H.E. and R.B. Tennent. 1977a. Site index equations for radiata pine in New Zealand. New Zealand Journal of Forestry Science 7(3): 408-416.

Burkhart, H.E. and R.B. Tennent. 1977b. Site index equations for Douglas-fir in Kaingaroa Forest. New Zealand Journal of Forestry Science 7(3): 417-419.

Byrne, J.C. and D.D. Reed. 1986. Complex compatible taper and volume estimation systems for red and loblolly pine. Forest Science 32(2): 423-443.

Candy, S.G. 1989. Compatible tree volume and variable-form stem taper models for *Pinus radiata* in Tasmania. New Zealand Journal of Forestry Science 19(1): 97-111.

Candy, S.G. 1997. Growth and yield models for *Eucalyptus nitens* plantations in Tasmania and New Zealand. Tasforests 9: 167-198.

Cao, Q.V. 2000. Prodiction of annual diameter growth and survival for individual trees from periodic measurements. Forest Science 46(1): 127-131.

Cao, Q.V. and V.C. Baldwin. 1999. A new algorithm for stand table projection models. Forest Science 45(4): 506-511.

Cao, Q.V., H.E. Burkhart and T.A. Max. 1980. Evaluation of two methods for cubic-volume prediction of loblolly pine to any merchantable limit. Forest Science 26(1): 71-80.

CIDE. 1967. Los suelos del Uruguay, su uso y manejo. Comision de Inversiones y Desarrollo Económico (CIDE), MGA, Montevideo, Uruguay.

Clutter, J.L. 1963. Compatible growth and yield models for loblolly pine. Forest Science 9: 354-371.

Clutter, J.L. and B.J. Allison 1974. A growth and yield model for *Pinus radiata* in New Zealand. In: J. Fries (Ed.). Growth models for tree and stand simulation. Stockholm. Royal Col. For. 136-160

Clutter, J.L. and F.A. Bennett. 1965. Diameter distribution in old-slash pine plantations. Georgia Forest Research Council. Report No. 13. 9 pp.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister and R.L. Bailey. 1983. Timber management: a quantitative approach. John Wiley & Sons, New York. 331 pp.

Clutter, J.L. and E.P. Jones. 1980. Prediction of growth after thinning in old-field slash pine plantations. USDA Forest Service. Research Paper. SE-217. 19 pp

Coile, T.S. 1948. Relation of soil characteristics to site index of loblolly and shotleaf pines in the lower Piedmont region of North Caroline. Duke University. School of Foresty Bulletin. 13.

Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. Forest Science 13(4): 365-375.

Daniels, R.F. and H.E. Burkhart. 1975. Simulation of individual tree growth and stand development in managed loblolly pine plantations. Division of Forest and Wildlife Resources, Virginia Polytechnic and Sate University. FWS-5-75.

Daniels, R.F. and H.E. Burkhart. 1988. An integrated system of forest stand models. Forest Ecology and Management 23: 159-177.

Demaerschalk, J.P. 1972. Converting volume equations to compatible taper equations. Forest Science 18(3): 241-245.

Demaerschalk, J.P. 1973. Integrated systems for the estimation of tree taper and volume. Canadian Journal of Forest Research 3: 90-94.

De Moraes Goncalves, J.L., J.L. Ioriatti Dematte and H.T. Zarate Do Couto. 1990. Relacoes entre a produtividade de sitios florestais de *Eucalyptus grandis* e *Eucalyptus saligna* com as propriedades de alguns solos de textura arenosa e media no estado de Sao Paulo (Relations between site productivity for *Eucalyptus grandis* and *Eucalyptus saligna* and some properties of sandy soils in Sao Paulo). IPEF Piracicaba No. 43-44: 24-39.

Dunlop, J. 1995. Permanent sample plot system - User manual. Forest Research Institute. FRI Bulletin. No. 187. 110 pp.

Dunningham, A.G. and M.E. Lawrence. 1987. An "interim" stand growth model for radiata pine grown in the Central North Island Pumice Plateau. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Echeverría, R.D. 2000. Estado de situación del *Eucalyptus globulus* en el Uruguay. In: Primer Seminario Internacional del Eucalyptus globulus en la Argentina. Mar del Plata. pp. 108-124

Ellis, J.C. and J.D. Dunlop. 1994. Minimum standards for collection of growth data from permanent sample plots. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Ellis, J.C. and J.D. Hayes. 1997. Field guide for sample plots in New Zealand forests. Forest Research Institute. FRI Bulletin. No. 186.

Fang, Z. and R.L. Bailey. 1999. Compatible volume and taper models with coefficients for tropical species on Hainan Island in Southern China. Forest Science 45(1): 85-100.

FAO. 1979. Eucalypts for planting. FAO Forestry series No. 11. FAO, Rome.

Figueiredo-Filho, A., B.E. Borders and K.L. Hitch. 1996. Taper equations for *Pinus taeda* plantations in Southern Brazil. Forest Ecology and Management 83: 39-46.

Franklin, E.C. and G. Meskimen 1983. Choice of species and provenances in cold summer climates. In: Proc. IUFRO/AFOCEL Frost Resistant Eucalypts Symposyim. Bordeaux, France. pp 341-357

Frazier, J.R. 1981. Compatible whole-stand and diameter distribution models for loblolly pine plantations. Ph.D. thesis. School of Forestry, Virginia Polytechnic Institute and State University, Blacksburg. 125 pp.

Furnival, G.M. 1961. An index for comparing equations used in constructing volume tables. Forest Science 7(4): 337-341.

Gadow, K.v. and G. Hui. 1999. Modelling forest development. Kluver Academic Publishers, Dordretch. 213 pp.

Gál, J. and I.E. Bella. 1995. Evaluation of stem taper functions for estimating log volume assortments. The Forestry Chronicle 71(6): 743-746.

Garcia de Leon, J.P. and A.R. Griffin 1995. Species/provenance evaluation and improvement strategy in the early stages of a eucalypt afforestation project in Uruguay. In: Eucalypt Plantations: Improving Fibre Yield and Quality. CRC-IUFRO Conference. Hobart. pp. 261-263

García, O. 1981. Simplified method-of-moments estimation for the Weibull distribution. New Zealand Journal of Forestry Science 11(3): 304-306.

García, O. 1984. New Class of growth models for even-aged stands: *Pinus radiata* in Golden Downs Forest. New Zealand Journal of Forestry Science 14(1): 65-88.

García, O. 1988a. Experience with an advanced growth modelling methodology. In: A. R. Ek, S. Shifley and T. E. Burk (Ed.). Forest growth modelling and prediction. Volume 2. USDA Forest Service. General Technical Report NC-120. p 668-675

García, O. 1988b. Growth modelling - a (Re)view. New Zealand Forestry 33(3): 14-17.

García, O. 1990. Growth of thinned and pruned stands. In: New approaches of spacing and thinning in plantation forestry. Rotorua. Forest Research Institute. FRI Bulletin No. 151. p: 84-97

García, O. 1991. What is a diameter distribution? In: Proceedings of the IUFRO Symposium on Integrated Forest Management Information Systems. Tsukuba, Japan.

García, O. 1992. Sampling for Tree Ring Analysis. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

García, O. 1994. The state-space approach in growth modelling. Canadian Journal of Forest Research 24: 1894-1903.

Gasana, J.K. and H. Loewenstein. 1984. Site classification for Midens's gum, *Eucalyptus globulus* subsp. *maidenii*, in Rwanda. Forest Ecology and Management 8: 107-116.

Glade, J.E. 1999. Curvas de índice de sitio para *Eucalyptus grandis* en Entre Ríos (Site index curves for *Eucalyptus grandis* in Entre Ríos). INTA, EEA Concordia. 10 pp.

Goelz, J.C.G. 2000. On the interrelatedness of all forest growth and yield models. In: C. J. Cieszewski (Ed.). First international conference on measurements and quantitative methods and management & the 1999 Southern Mensurationists Meeting. Jekyll Island, Georgia. D.B. Warnell School of Forest Resources. University of Georgia. 44-65

Goodwin, A.N. and S.G. Candy. 1986. Single-tree and stand growth models for a plantation of *Eucalyptus globulus* Labill. in northern Tasmania. Australian Forest Research 16(2): 131-144.

Gordon, A. 1983. Estimating bark thickness of *Pinus radiata*. New Zealand Journal of Forestry Science 13(3): 340-353.

Gordon, A.D. 1996. Projecting inventory data: predicting individual tree height growth. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Gordon, A.D. and M.E. Lawrence. 1992. A Comparison of methods to predict individual tree diameter growth. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Gordon, A.D. and M.E. Lawrence. 1994. Projecting inventory data: predicting individual tree diameter growth. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Gordon, A.D. and R.G. Shula. 1999. A Guide to using the individual-tree growth model. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Gordon, A.D., C. Lundgren and E. Hay. 1995. Development of a composite taper equation to predict over- and under-bark diameter and volume of *Eucalyptus saligna* in New Zealand. New Zealand Journal of Forestry Science 25(3): 318-327.

Gordon, A.D., C. Lundgren and E. Hay. 1999. Composite taper equations to predict over- and under-bark diameter and volume of *Eucalyptus pilularis*, *E. globoidea* and *E. muelleriana* in New Zealand. New Zealand Journal of Forestry Science 29(2): 311-317.

Goulding, C.J. and J.C. Murray. 1976. Polynomial taper equations that are compatible with tree volume equations. New Zealand Journal of Forestry Science 5(3): 313-322.

Goulding, C.J. 1994. Development of growth models for *Pinus radiata* in New Zealand - experience with management and process models. Forest Ecology and Management 69: 331-343.

Goulding, C.J. (1995). Growth and yield models. In: D. Hammond (Ed.). Forestry Handbook (3rd. Edition). Christchurch, N.Z. Institute of Forestry Inc. pp 111-114.

Goulding, C.J. and J.W. Shirley 1979. A method to predict the yield of log assortments for long term planning. In: D. A. Elliot (Ed.). Mensuration for management planning of exotic forest plantations. Forest Research Institute, Symposium No. 20. 301-314

Gujarat, D. 1970. Use of dummy variables in tensting for equality between sets of coefficients in linear regression: a generalization. American Statistician 25(4): 21-33.

Hafley, W.L. and H.T. Schreuder. 1977. Statistical distribution for fitting diameter and height data in even-aged stands. Canadian Journal of Forest Research 7: 481-487.

Hamilton, D.A. 1974. Event probabilities estimated by regression. USDA Forest Service. Research Paper. INT-152. 18 pp.

Hamilton, D.A. and B.M. Edwards. 1976. Modeling the probability of individual tree mortality. USDA Forest Service. Research Paper. INT-85. 22 pp.

Hann, D.W. and D.R. Larsen. 1991. Diameter growth equations for fourteen tree species in southwest Oregon. Forest Research Laboratory, Oregon State University. Research Bulletin. No. 69.

Harrison, W.C. and R.F. Daniels 1988. A new biomatematical model for growth and yield of loblolly pine plantations. In: S. R. Ek, S. R. Shifley and T. E. Burk (Ed.). Proceedings of the IUFRO Conference on Forest Growth modelling and Prediction. Mineapolis, MN. USDA Forest Service, General Technical Report NC-120. 293-304

Hayward, W.J. 1987. Volume and taper of *Eucalyptus regnans* grown in the central North Island of New Zealand. New Zealand Journal of Forestry Science 17(1): 109-120.

Hayward, W., I. Jenkin and D. New. 1987. Evaluation of the performance of growth models for Kaingaroa. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Hilt, D.E. 1980. Taper-based system for estimating stem volumes for upland oak. USDA Forest Servive. Research Paper. NE-458. 12 pp

Huang, S., D. Price and S.J. Titus. 2000. Development of ecoregion-based height-diameter models for white spruce in boreal forests. Forest Ecology and Management 129: 125-141.

Huang, S.M., S.J. Titus and D.P. Wiens. 1992. Comparison of non-linear height-diameter functions for major Alberta tree species. Canadian Journal of Forest Research 22(9): 1297-1304.

Huebschmann, M.M., L.R. Gering, T.B. Lynch, O. Bitoki and P.A. Murphy. 2000. An individual-tree growth and yield prediction system for uneven-aged shortleaf pine stands. Southern Journal of Applied Forestry 24(2): 112-121.

Hunter, I.R. and A.R. Gibson. 1984. Predicting *Pinus radiata* site index from environmental variables. New Zealand Journal of Forestry Science 14(1): 53-64.

Husch, B., C.I. Miller and T.W. Beers. 1972. Forest mensuration. 2nd Edition. Ronald Press, New York. 410 pp.

Hyink, D.M. and J.W. Moser. 1983. A generalized framework for projecting forest yield and stand structure using diameter distributions. Forest Science 29(1): 85-95.

Jerram, M.R.K. 1939. Elementary Forest Mensuration. Thomas Murby & Co., London. 124 pp.

Knoebel, B.R., H.E. Burkhart and D.E. Beck. 1986. A growth and yield model for thinned stands of yellow-poplar. Forest Science Monograph 27: 62 pp.

Knowe, S.A. 1994. Incorporating the effects of interespecific competition and vegetation-management treatments in stand table projection models for Douglas-fir saplings. Forest Ecology and Management 67: 87-99.

Knowe, S.A. and W.I. Stein. 1995. Predicting the effects of site preparation on development of young Douglas-fir plantations. Canadian Journal of Forest Research 25: 1538-1547.

Knowe, S.A., G.R. Ahrens and D.S. DeBell. 1997. Comparison of diameter-distribution-prediction, stand-table-projection, and individual-tree-growth modeling approaches for young red alder plantations. Forest Ecology and Management 98: 49-60.

Kozak, A. 1988. A variable-exponent taper equation. Canadian Journal of Forest Research 18: 1363-1368.

Kozak, A. 1997. Effects of multicollinearity and autocorrelation on the variable-exponent taper functions. Canadian Journal of Forest Research 27: 619-629.

Kozak, A. and J.H.G. Smith. 1993. Standards for evaluating taper estimation systems. The Forestry Chronicle 69(4): 438-444.

Kuru, G.A. 1989. A Diameter distribution growth and yield model for naturally regenerated and seeded stands of radiata pine in Kaingaroa Forest. M.For.Sc. thesis. School of Forestry, University of Canterbury, Christchurch. 101 pp.

Kuru, G.A., A.G.D. Whyte and R.C. Woollons. 1992. Utility of reverse Weibull and extreme value density functions to refine diameter distribution growth estimates. Forest Ecology and Management 48: 165-174.

Landsberg, J.J. and R.H. Waring. 1997. A generalised model of forest productivity using simplified concepts of radiation-use efficiency, carbon balance and partitioning. Forest Ecology and Management 95: 209-228.

Law, K.R.N. 1990. A growth model for Douglas-fir grown in the South Island of New Zealand. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Lawrence, M.E. 1990. Diameter distributions for the regional stand growth models. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Lee, S.H. 1998. Modelling growth and yield of Douglas-fir using different interval lengths in the South Island of New Zealand. Ph.D. thesis. School of Forestry, University of Canterbury, Christchurch. 239 pp.

Lenhart, J.D. and J.L. Clutter. 1971. Cubic-foot yield tables for old-field loblolly pine plantations in the Georgia Piedmont. Georgia Forest Research Council. Report No. 22. 13 pp.

Lindsay, S.R., G.R. Wood and R.C. Woollons. 1996. Stand table modelling through the Weibull distribution and usage of skewness information. Forest Ecology and Management 81: 19-23.

Louw, J.H. 1997. A site-growth study of Eucalyptus grandis in the Mpumalanga Escarpment Area. South African Forestry Journal 180: 1-14.

Lundgren, C. and A.D. Gordon. 1997. Modelling individual tree survival of radiata pine in New Zealand. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Lynch, T.B., K.L. Hitch, M.M. Huebschmann and P.A. Murphy. 1999. An individual-tree growth and yield prediction system for even-aged natural shortleaf pine forests. Southern Journal of Applied Forestry 23(4): 203-211.

Maclaren, J.P., J.C. Grace, M.O. Kimberley, R.L. Knowles and G.G. West. 1995. Height growth of *Pinus radiata* as affected by stocking. New Zealand Journal of Forestry Science 25(1): 73-90.

Manley, B.R. 1981. A distance-independent tree growth model for radiata pine in New Zealand. Ph.D. thesis. University of Washington. 106 pp.

Martin, S.W., R.L. Bailey and E.J. Jokela. 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. Southern Journal of Applied Forestry 23(1): 39-45.

Mason, E.G. 1992. Decision-support systems for establishing radiata pine plantations in the Central North Island of New Zealand. Ph.D. thesis. School of Forestry, University of Canterbury, Christchurch. 251 pp.

Mason, E.G. 2001. A model of the juvenile growth and survival of *Pinus radiata* D. Don: Adding the effects of initial seedling diameter and plant handling. New Forests (in press).

Mason, E.G. and A.G.D. Whyte. 1997. Modelling initial survival and growth of radiata pine in New Zealand. Acta Forestalia Fennica 255: 38 pp.

Mason, E.G., A.G.D. Whyte, R.C. Woollons and B. Richardson. 1997. A model of the growth of juvenile radiata pine in the Central North Island of New Zealand: links with older models and rotation-length analyses of the effects of site preparation. Forest Ecology and Management 97: 187-195.

Matney, T.G. and A.D. Sullivan 1982. Approximating thinned stand diameter distribution with statistical probability functions. In: P. Jones (Ed.). Proceedings of the second biennial Southern Silviculture Research Conference. USDA Forest Service SE For. Experimental Station.

Max, T.A. and H.E. Burkhart. 1976. Segmented polynomial regression applied to taper equations. Forest Science 22(3): 283-289.

Mayer, D.G. and D.G. Butler. 1993. Statistical validation. Ecological Modelling 68: 21-32.

McClure, J.P. and R.L. Czaplewski. 1986. Compatible taper equation for loblolly pine. Canadian Journal of Forest Research 16: 1272-1277.

McDill, M. 2000. A comparison of two growth models developed for regional timber inventory projection. In: Seventh Symposium on systems analysis in forest resources. Traverse, Michigan. USDA Forest Service, North Central Research Station. General Technical Report NC-205. 330-336

McEwen, A.D. 1978. Forest Service computer system for permanent sample plots. In: Mensuration for management planning of exotic forest plantations. FRI Symposium No. 20. p 235-252

McGee, C.E. and L. Della-Bianca. 1967. Diameter distributions in natural yellow-poplar stands. USDA Forest Service. Research Paper. SE-25. 7 pp.

Monsreud, R.A. 1976. Simulation of forest tree mortality. Forest Science 22(4): 438-444.

Mountfort, C.J. 1978. Growth of Douglas fir. New Zealand Forest Service. FRI Symposium No. 15. pp 375-411.

Mountfort, C.J. 1979. *P. radiata* site index/altitudinal relationship for forests established on pumice soil. In: D. A. Elliot (Ed.). Mensuration for forest management planning of exotic plantations, Proceedings of FRI Symposium No. 20. New Zealand. Forest Research Institute. 217-226

Muhairwe, C.K. 1999. Taper equations for *Eucalyptus pilularis* and *Eucalyptus grandis* for the north coast in New South Wales, Australia. Forest Ecology and Management 113: 251-269.

Munro, D.D. 1974. Forest growth models - a prognosis. In: J. Fries (Ed.). Growth models for tree and stand simulation. Stockholm, Sweden. Royal College of Forestry.

Munro, D.D. 1984. Growth modelling for fast-growing plantations - a Review. In: D. C. Grey, A. P. G. Schonau and C. J. Schutz. (Ed.). Symposium on Site and Productivity of Fast Growing Plantations. Pretoria, South Africa. IUFRO. Vol 1. 333-344

Murphy, P.A. and R.M. Farrar. 1988. Basal-area projection equations for thinned natural even-aged forest stands. Canadian Journal of Forest Research 18: 827-832.

Murphy, P.A. and D.L. Graney. 1998. Individual-tree basal area growth, survival and total height models for upland hardwoods in the Boston Mountains of Arkansas. Southern Journal of Applied Forestry 22(3): 184-192.

Murphy, P.A. and M.G. Shelton. 1996. An individual-tree basal area growth model for loblolly pine stands. Canadian Journal of Forest Research 26: 327-331.

Myers, R.H. 1990. Classical and modern regression with applications. PWS-KENT Publishing Company, Boston. 488 pp.

Nelson, T.C. 1964. Diameter distribution and growth of loblolly pine. Forest Science 10: 105-115.

Nepal, S.K. and G.L. Somers. 1992. A generalized approach to stand table projection. Forest Science 38(1): 120-133.

Newberry, J.D. and H.E. Burkhart. 1986. Variable-form stem profile models for loblolly pine. Canadian Journal of Forest Research 16: 109-114.

Newnham, R.M. 1992. Variable-form taper functions for four Alberta tree species. Canadian Journal of Forest Research 22: 210-223.

Noble, A.D., M.J. Donkin and C.W. Smith 1991. The importance of Soil Properties as Indicators of Site Quality for *Eucalyptus grandis* on the Zululand Coastal Plain. In. IUFRO Proceedings. Volume I. 433-443

Olson, D.F. and L. Della-Bianca. 1959. Site index comparisons for several tree species in the Virginia-Carolina Piedmont. USDA Forest Service. S.E. Forest Experimental Station Paper 104.

Ormerod, D.W. 1973. A simple bole model. The Forestry Chronicle 49(3): 136-138.

Penman, J.T.D. 1988. Volume, taper and bark thickness in seedlings and cuttings from Mamaku Forest, New Zealand. New Zealand Journal of Forestry Science 18(3): 311-317.

Perez, D.N., H.E. Burkhart and C.T. Stiff. 1990. A variable-form taper function for *Pinus oocarpa* Schiede in Central Honduras. Forest Science 36(1): 186-191.

Pienaar, L.V. 1989. A stand table projection approach to yield prediction in plantations. South African Forestry Journal 149: 44-47.

Pienaar, L.V. and W.W. Harrison. 1988. A stand table projection approach to yield prediction in unthinned even-aged plantations. Forest Science 34(3): 804-808.

Pienaar, L.V. and J.W. Rheney. 1995. Modeling stand level growth and yield response to silvicultural treatments. Forest Science 41(3): 629-639.

Pienaar, L.V. and K.J. Turnbull. 1973. The Chapman-Richards generalization of Von Bertalanffy's growth model for basal area growth and yield in even-aged stands. Forest Science 19(1): 2-22.

Prodan, M., R. Peters, F. Cox and P. Real. 1997. Mensura Forestal. Serie Investigación y Educación en Desarrollo Sustentable. IICA - BMZ/GTZ. 561 pp.

Rauscher, H.M., M.J. Young, C.D. Webb and D.J. Robinson. 2000. Testing the accuracy of growth and yield models for Southern hardwood forests. Southern Journal of Applied Forestry 24(3): 176-185.

Rawlings, J.O. 1988. Applied regression analysis - A research tool. Wadworth, Belmont. 553 pp.

Real, P.L. and J.A. Moore. 1988. An individual-tree taper system for Douglas-fir in the inland Northwest. In: Ek, A.R., Shifley, S.K. and Burk, T.E. (Eds.). 1988. Forest Growth Modelling and Prediction. Proceedings of IUFRO Conference, August 24-28, 1987, Minneapolis, MN. USDA For. Serv. Gen. Tech. Rep. NC-120.

Reed, D.D. and J.C. Byrne. 1985. A simple, variable form volume estimation system. The Forestry Chronicle 61(2): 87-90.

Reineke, L.H. 1933. Perfecting a stand-density index for even-aged forests. Journal of Agricultural Research 46(7): 627-638.

Reynolds, M.R., T.E. Burk and W. Huang. 1988. Goodnes-of-fit tests and model selection procedures for diameter distribution models. Forest Science 34(2): 373-399.

Richards, F.J. 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10: 290-300.

Ritchie, M.W. and D.W. Hann. 1997. Evaluation of individual-tree and dissagregative prediction methods for Douglas-fir stands in western Oregon. Canadian Journal of Forest Research 27: 207-216.

Ritchie, M.W. and D.W. Hann. 1997. Implications of dissagregation in forest growth and yield modelling. Forest Science 43(2): 223-233.

Rodrigues da Cunha Neto, F., J.R. Soares Scolforo, N. Calegario, A. Donizette de Oliveira and H. Kanegae Júnior. 1994. Modelo para predicao da producao por classe de diametro para *Eucalyptus grandis* (Model to predict production per diameter-class for *E. grandis*). Cerne 1(1): 108-122.

SAS. Institute Inc. 1989. SAS/STAT User's guide, Version 6, Fourth Edition, Volume 2. SAS Institute Inc., Cary, NC. 846 pp.

SAS. Institute Inc. 1990. SAS Procedures guide, Version 6. SAS Institute Inc., Cary. 705 pp.

SAS. Institute Inc. 1996. SAS/STAT Software: Changes and Enhancements through Release 6.11. SAS Institute Inc., Cary, NC. 1104 pp.

Schreuder, H.T., W.L. Hafley and F.A. Bennett. 1979. Yield predictions for unthinned natural slash pine stands. Forest Science 25: 25-30.

Schumacher, F.X. 1939. A new growth curve and its application to timber yield studies. Journal of Forestry 37: 819-820.

Shifley, S.R. and G.J. Brand. 1984. Chapman-Richards growth function constrained for maximum tree size. Forest Science 30(4): 1066-1070.

Shifley, S. and S. Fairweather 1983. Individual-tree diameter growth and mortality models for Western Oregon. In: J. F. Bell and T. Atterbury (Ed.). Renewable resource inventories for monitoring changes and trends. Corvallis. Oregon State University. 737 pp

Shiver, B.D. 1988. Sample sizes and estimation methods for the Weibull distribution for unthinned slash pine plantation diameter distributions. Forest Science 34(3): 809-814.

Shortt, J.S. and H.E. Burkhart. 1996. A comparison of loblolly pine plantation growth and yield models for inventory updating. Southern Journal of Applied Forestry 20(1): 15-22.

Shula, R.G. 1997a. Modelling individual-tree probability of survival of unthinned radiata pine. Unpublished Report of Forest Research's Stand Growth Modelling Co-operative.

Shula, R.G. 1997b. Projecting inventory data: revised individual-tree diameter growth equations. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Shula, R.G. 1997c. Projecting inventory data: revised individual-tree static height equation. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Smalley, G.W. and R.L. Bailey. 1974. Yield tables and stand structure for loblolly pine plantations in Tennessee, Alabama and Georgia highlands. USDA Forest Service. Sth For. Exp. Stn. Research Paper. SO-96.

Smith, W.B. 1983. Adjusting STEMS regional forest growth model to improve local predictions. USDA Forest Service. Research Note. NC-197. 5 pp.

Snowdon, P. 2000. Annual updating of plantation inventory estimates using hybrid models. In: M. Hansen and T. Burl (Ed.). Integrated Tools for Natural Resources Inventories in the 21st century. 562-569

Snowdon, P., T. Jovanovic and T.H. Booth. 1999. Incorporation of indices of annual climatic variation into growth models of *Pinus radiata*. Forest Ecology and Management 117: 187-197.

Snowdon, P., R.C. Woollons and M.L. Benson. 1998. Incorporation of climatic factors into models of growth of *Pinus radiata* in a spacing experiment. New Forests 16: 101-123.

Soares, P., M. Tomé, J.P. Skovsgaard and J.K. Vanclay. 1995. Evaluating a growth model for forest management using continuous forest inventory data. Forest Ecology and Management 71: 251-265.

Somers, G.L. and S.K. Nepal. 1994. Linking individual-tree and stand-level growth models. Forest Ecology and Management 69: 233-243.

Sorrentino, A. 1991. Indices de sitio preliminares para las principales especies forestales, cultivadas en el Uruguay (Provisional site indices for the main forest species cultivated in Uruguay). Facultad de Agronomia, Universidad de la Republica. Boletin de Investigacion. No. 33.

Sorrentino, A. 1992. Proyecto Indices de Sitio, Volumetría y crecimiento de pinos y eucaliptos en el Uruguay (Project: Site indices, volume yields and growth of pines and eucalypts in Uruguay). Unpublished report.

Sorrentino, A. 1993. Comportamiento de Especies Exoticas. Parte I: Indices de Sitios (Performance of exotic species. Part I: Site Indices). In: Congreso de Ingenieros Agronomos. Montevideo, Uruguay. 15-19

Stage, A.R. 1973. Prognosis model for stand development. USDA Forest Service. Research Paper. INT-137.

Staudhammer, C. and V. LeMay. 2000. Height prediction equations using diameter and stand density measures. The Forestry Chronicle 76(2): 303-309.

Sullivan, A.D. and J.L. Clutter. 1972. A simultaneous growth and yield model for loblolly pine. Forest Science 18: 76-86.

Tassisa, G. and H.E. Burkhart. 1998. An application of mixed effects analysis to modeling thinning effects on stem profile of loblolly pine. Forest Ecology and Management 103: 87-101.

Temu, M., J. 1992. Forecasting yield of Douglas-fir in the South Island of New Zealand. Ph.D. thesis. School of Forestry, University of Canterbury, Christchurch. 210 pp.

Tennent, R.B. 1981. A simulation model for *Pinus radiata*. Ph.D. dissertation. University of Georgia. pp.

Thomas, C.E. and B.R. Parresol. 1991. Simple, flexible, trigonometric taper equations. Canadian Journal of Forest Research 21(7): 1132-1137.

Ure, J. 1950. The natural vegetation of the Kaingaroa Plains as an indicator of site quality for exotic conifers. New Zealand Journal of Forestry 6: 112-123.

Vanclay, J.K. 1994. Modelling forest growth and yield - Applications to mixed tropical forests. CAB International, Wallingford. 312 pp.

Vanclay, J.K., J.P. Skovsgaard and C. Pilegaard Hansen. 1995. Assesing the quality of permanent sample plot databases for growth modelling in forest plantations. Forest Ecology and Management 71: 177-186.

Villanueva, T. 1992. Integrated yield forecasting and harvest scheduling in a tropical pine plantation in Fiji. PhD. thesis. School of Forestry, University of Canterbury, Christchurch. 208 pp.

Von Bertalanffy, L. 1957. Quantitative laws in metabolism and growth. Quarterly Review of Biology 32: 317-231.

Walsh, T.A. 1986. Diameter/basal area increment equations for loblolly pine trees in cutover, site-prepared plantations. M.S. thesis. School of For. and Wildl. Res., Va. Polytech. Inst. and State Univ, Blacksburg, VA. 74 pp.

West, G.G. 1997. The effect of Maku Lotus on the growth of radiata pine in Kaingaroa Forest. M.For.Sc. thesis. School of Forestry, University of Canterbury, Christchurch. 195 pp.

West, G.G., N.J. Eggleston and J. McLanachan. 1987. Further developments and validation of the Early growth model. Forest Research Institute, Rotorua. 32 pp.

West, G.G., R.L. Knowles and A.R. Koehler. 1982. Model to predict the effects of pruning and early thinning on the growth of radiata pine. Forest Research Institute, Rotorua. 35 pp.

West, P.W. 1980. Use of diameter increment and basal area increment in tree growth studies. Canadian Journal of Forest Research 10: 71-77.

West, P.W. 1995. Application of regression analysis to inventory data with measurements on succesive occasions. Forest Ecology and Management 71: 227-234.

Whyte, A.G.D., M. Temu, J. and R.C. Woollons 1992. Improving yield forecasting reliability through aggregated modelling. In: G. B. Wood and B. J. Turner (Ed.). Integrating forest information over space and time. Camberra. The Australian National University. IUFRO S4.01 and S4.02 Conference Proceedings. 81-88

Whyte, A.G.D. and R.C. Woollons 1992. Diameter distribution growth and yield modelling: recent revisions and perspectives. In: G. B. Wood and B. J. Turner (Ed.). Integrating forest information over space and time. Camberra. The Australian National University. IUFRO S4.01 and S4.02 Conference Proceedings. 89-93

Wilcox, P.L. 1987. The retrospective sampling technique. Unpublished report of Forest Research's Stand Growth Modelling Co-operative.

Woollons, R.C. 1998. Even-aged stand mortality estimation through a two-step regression process. Forest Ecology and Management 105: 189-195.

Woollons, R.C. and W.J. Hayward. 1985. Revision of a growth and yield model for radiata pine in New Zealand. Forest Ecology and Management 11: 191-202.

Woollons, R.C., R. Sands and P. Snowdon. 1998. Influence of climate on top height and diameter development in Pinus radiata forests in the Hawke's Bay Region of New Zealand. Commonwealth Forestry Review 77(4): 267-271.

Woollons, R.C., A.G.D. Whyte and L. Xu. 1990. The Hossfeld function: an alternative model for depicting stand growth and yield. Journal of the Japanese Forestry Society 15: 25-35.

Woollons, R.C. and G.B. Wood 1992. Utility and performance of five sigmoid yield-age functions, fitted to stand growth data. In: G. B. Wood and B. J. Turner (Ed.). Integrating forest information over space and time. Camberra. The Australian National University. IUFRO S4.01 and S4.02 Conference Proceedings. 71-80

Xu, L. 1990. Growth and Yield of Douglas-fir plantations in the Central North Island of New Zealand. Ph.D. thesis. School of Forestry, University of Canterbury, Christchurch. 244 pp.

Xu, L., G.R. Wood, R.C. Woollons and A.G.D. Whyte. 1992. Stand table prediction with Reverse Weibull and Extreme Value density functions: some theoretical considerations. Forest Ecology and Management 48: 175-178.

Yang, R.C., A. Kozak and J.H.G. Smith. 1978. The potential of Weibull-type functions as flexible growth curves. Canadian Journal of Forest Research 8(4): 424-431.

Zhang, L., J.A. Moore and J.D. Newberry. 1993. Dissagregating stand volume growth to individual trees. Forest Science 39(2): 295-308.

Zhang, S., R.L. Amateis and H.E. Burkhart. 1997a. Constraining individual tree diameter increment and survival models for loblolly pine plantations. Forest Science 43(3): 414-423.

Zhang, S., H.E. Burkhart and R.L. Amateis. 1996. Modeling individual tree growth for juvenile loblolly pine plantations. Forest Ecology and Management 89: 157-172.

Zhang, S., H.E. Burkhart and R.L. Amateis. 1997b. The influence of thinning on tree height and diameter relationships in loblolly pine plantations. Southern Journal of Applied Forestry 21(4): 199-205.

Zhao, W. 1999. Growth and Yield Modelling of *Pinus radiata* in Canterbury, New Zealand. Ph.D. thesis. School of Forestry, University of Canterbury, Christchurch. 192 pp.

APPENDICES

Appendix 1 - Difference equations assayed

Chapman-Richards anamorphic $$Y_2 = Y_1\left[\frac{1-\exp(-bt_2)}{1-\exp(-bt_1)}\right]^c$$

Chapman-Richards polymorphic (1) $$Y_2 = a\left(\frac{Y_1}{a}\right)^{\frac{\ln(1-\exp(-kt_2))}{\ln(1-\exp(-kt_1))}}$$

Chapman-Richards polymorphic (2) $$Y_2 = a\left\{1-\left[1-\left(\frac{Y_1}{a}\right)^{1-m}\right]^{\frac{t_2}{t_1}}\right\}^{\frac{1}{1-m}}$$

Chapman-Richards polymorphic (3) $$Y_2 = a\left\{1+\left[\left(\frac{a}{Y_1}\right)^{v}-1\right]\exp[-b(t_2-t_1)]\right\}^{-\frac{1}{v}}$$

Gompertz anamorphic $$Y_2 = Y_1\frac{\exp[-b\exp(-ct_2)]}{\exp[-b\exp(-ct_1)]}$$

Gompertz polymorphic (1) $$Y_2 = \exp\{\ln(Y_1)\exp[-b(t_2-t_1)]\}\exp\{a[1-\exp(-b(t_2-t_1))]\}$$

Gompertz polymorphic (2)

$$Y_2 = \exp\left\{\ln(Y_1)\exp\left[-b(t_2-t_1)+c\left(t_2^2-t_1^2\right)\right]\right\}\exp\left\{a\left[1-\exp\left(-b(t_2-t_1)+c\left(t_2^2-t_1^2\right)\right)\right]\right\}$$

Hossfeld anamorphic $$Y_2 = \frac{1}{\left(\frac{1}{Y_1}\right)+c\left(\frac{1}{t_2^b}-\frac{1}{t_1^b}\right)}$$

Hossfeld polymorphic $$Y_2=\left\{\left(\frac{1}{Y_1}\right)\left(\frac{t_1}{t_2}\right)^b+a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\}^{-1}$$

Levakovic anamorphic $$Y_2=Y_1\left[\left(\frac{t_2}{t_1}\right)^2\left(\frac{b+t_1^2}{b+t_2^2}\right)\right]^c$$

Levakovic polymorphic $$Y_2=\left\{Y_1^c\left(\frac{t_1}{t_2}\right)^2+a\left[1-\left(\frac{t_1}{t_2}\right)^2\right]\right\}^{\frac{1}{c}}$$

Monomolecular anamorphic $$Y_2=Y_1\left[\frac{1-b\exp(-ct_2)}{1-b\exp(-ct_1)}\right]$$

Monomolecular polymorphic $$Y_2=Y_1\exp[-b(t_2-t_1)]+a\{1-\exp[-b(t_2-t_1)]\}$$

Schumacher anamorphic $$Y_2=Y_1\exp\left[-b\left(\frac{1}{t_2^c}-\frac{1}{t_1^c}\right)\right]$$

Schumacher polymorphic (1) $$Y_2=\exp\left\{\ln(Y_1)\left(\frac{t_1}{t_2}\right)+a\left[1-\left(\frac{t_1}{t_2}\right)\right]\right\}$$

Schumacher polymorphic (2) $$Y_2=\exp\left\{\ln(Y_1)\left(\frac{t_1}{t_2}\right)^b+a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\}$$

Weibull anamorphic $$Y_2=Y_1\frac{\left[1-\exp\left(-bt_2^c\right)\right]}{\left[1-\exp\left(-bt_1^c\right)\right]}$$

Weibull polymorphic $$Y_2=a-b\left[\frac{(a-Y_1)}{b}\right]^{\left[\left(\frac{t_2}{t_1}\right)^c\right]}$$

Appendix 2 - Basal area model for *P. menziesii* with dummy variables for *Phaeocryptopus*-free (prior to 1966) stands

$$G_2 = G_1^{(t_1/t_2)^b} \exp\left\{a\left[1-\left(\frac{t_1}{t_2}\right)^b\right]\right\}$$

where $a=3.7756+2.633\left(\frac{ALT}{1000}\right)+\frac{5.2259}{tt}\left(\frac{G_t}{G_b}\right)$

$$b=1.7406-1.7491\left(\frac{ALT}{1000}\right)+0.2225D$$

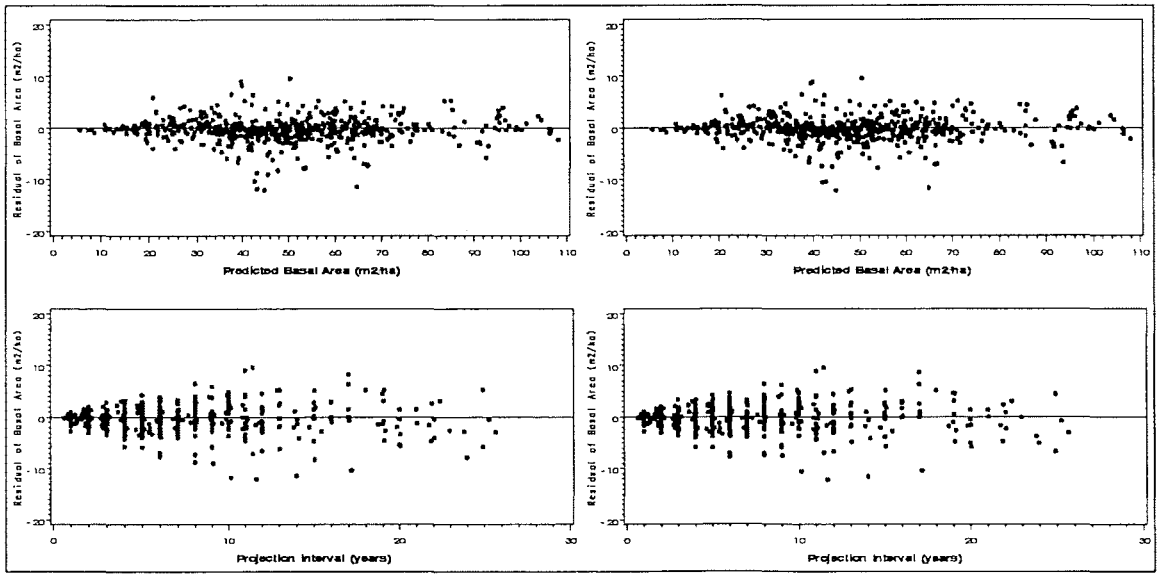
G_t and G_b are stand basal areas thinned and before last thinning, respectively

ALT: altitude in metres above sea level

D=1 if measurement taken prior to or in 1966; D=0 otherwise

Basal area estimations for the period after 1966 obtained from this model differ to estimations from Equation 9.2 with parameters in Table 9.3. Using a random subsample of 4 measurements per plot (n=625), the following residuals were obtained from both models:

	Model with dummy variable	Selected model (fitted to post 1966 data)
Mean residual	-0.28	-0.19
Standard deviation	2.34	2.30



Appendix 3 - Diameter potential index of a tree (DPIT) and change in potential diameter (CHG_PD). Adapted from Shula (1997a, 1997b, 1997c).

Selected difference equations:

E. grandis Monomolecular polym. $dbh_2 = dbh_1 \exp[-b(t_2 - t_1)] + a\{1 - \exp[-b(t_2 - t_1)]\}$

P. menziesii Schumacher anamorphic $dbh_2 = dbh_1 \exp\left[-b\left(\frac{1}{t_2^c} - \frac{1}{t_1^c}\right)\right]$

where $b = b_0 + b_1 * \ln(ALT)$ and $c = c_0 + c_1 * \ln(ALT)$

P. radiata Weibull polymorphic $dbh_2 = a - b\left[\frac{(a - dbh_1)}{b}\right]^{\left[\left(\frac{t_2}{t_1}\right)^c\right]}$

Procedure:

(i) Selected equations were fitted to mean diameters from the tallest 75 trees per hectare (equivalent to 3 trees in a 0.04-ha plot) and the following parameter estimates were obtained:

Species	a	b / b_0	c / c_0	b_1	c_1	Fit Index
E. grandis	42.9884	0.1116	-	-	-	0.934
P. menziesii	-	-13.0597	-0.8712	3.5461	0.2263	0.960
P. radiata	166.9569	197.88578	0.5343115	-	-	0.942

(ii) Selected equations with parameters in the above table were applied to each tree with t_2 made equal to the base age (10, 20 and 40 years for *E. grandis*, *P. radiata* and *P. menziesii*, respectively). Predicted diameters at base age represent the diameter potential index of a tree (DPIT)

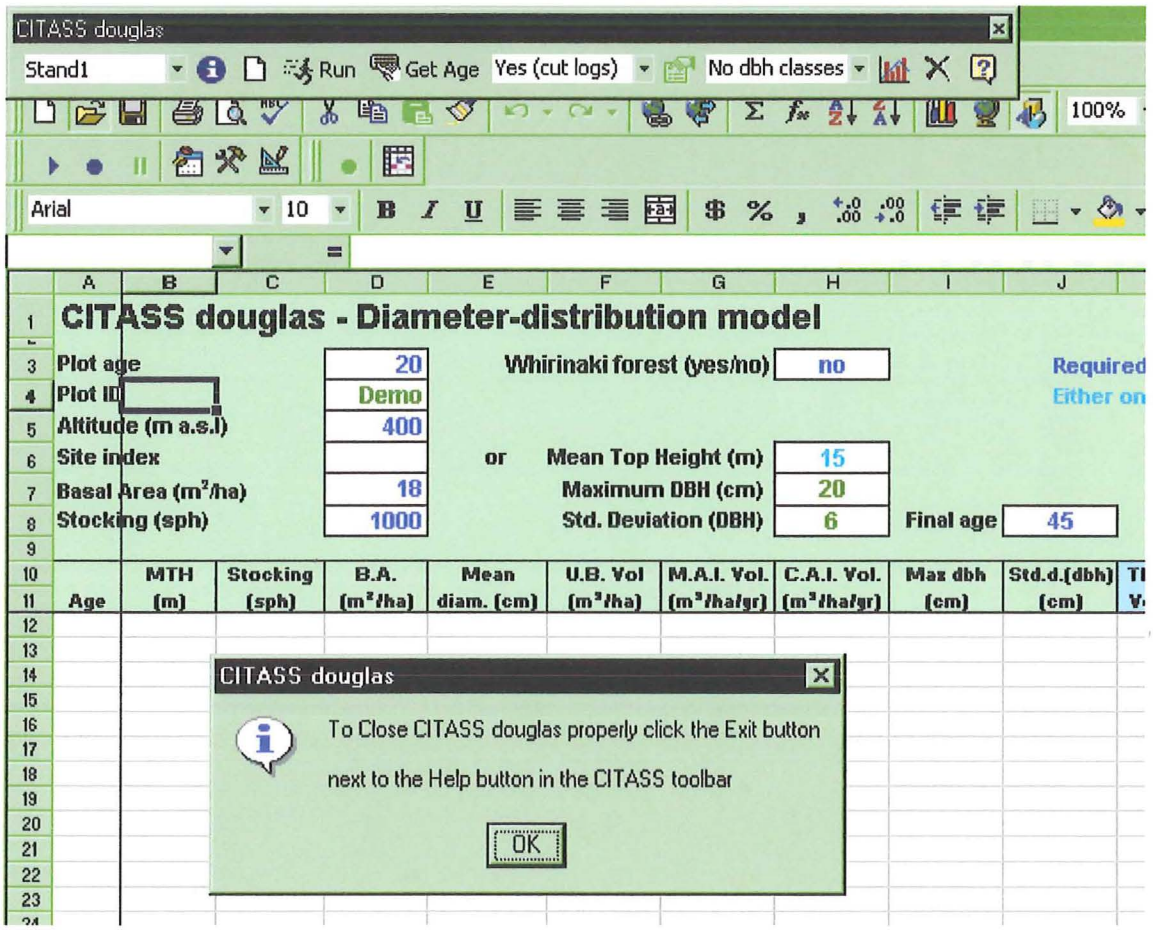
(iii) DPIT was substituted for dbh_1, then ages were set accordingly (i.e. t_1=base age and t_2=age at the end of the interval). Predicted diameters so obtained were considered potential diameters of each tree (PDT_i) at age t_2.

(iv) Diameters at the beginning of the period (dbh_1) were subtracted from PDT_i and the difference was divided by the projection interval (to accommodate for projection intervals that were not exactly one year). This procedure estimated the change in potential diameter (CHG_PD).

Appendix 4 Compatible Individual-Tree And Stand Simulators (CITASS) for *E. grandis*, *P. menziesii* and *P. radiata*.

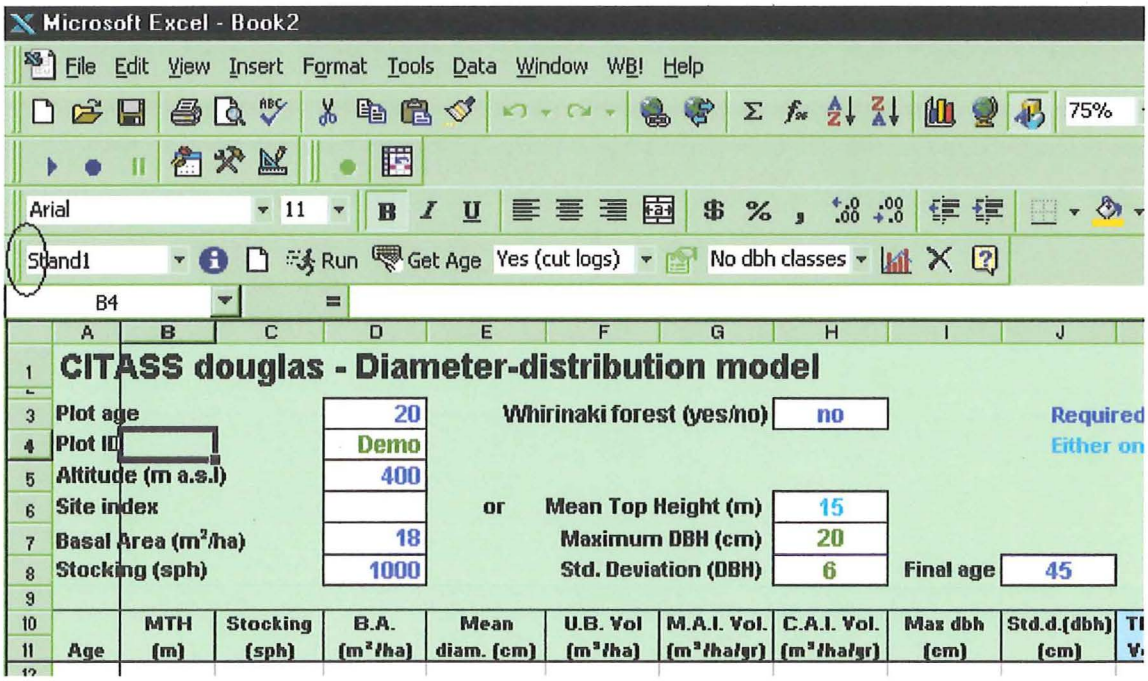
The attached diskette contains 3 Excel add-in files called "CITASS grandis.xla", "CITASS douglas.xla" and "CITASS radiata.xla". The following paragraphs briefly describe CITASS douglas (as an example, the other two programs are almost identical).

1) Copy the XLA file to your hard drive.
2) Open the XLA file from Microsofrt Excel 1997 or later versions (enable macros).
3) An Excel file with tree sheets will be created as well as the CITASS toolbar (see picture below). The new book is not the XLA file (which is hidden and cannot be modified) but a normal XLS file (e.g. Book1.xls) that you can save in any folder.



4) The CITASS toolbar can be located with other Excel toolbars by clicking its title bar and dropping it in the desired position. Once the CITASS toolbar is with the other

toolbars it can be moved by clicking the two pale vertical lines on the left (circled area in the picture below).



5) The main features of the program are described in the 2-page online help (see below) that can be activated by clicking the Help button on the right-hand side of the CITASS toolbar.

CITASS douglas - Help Page 1 of 2

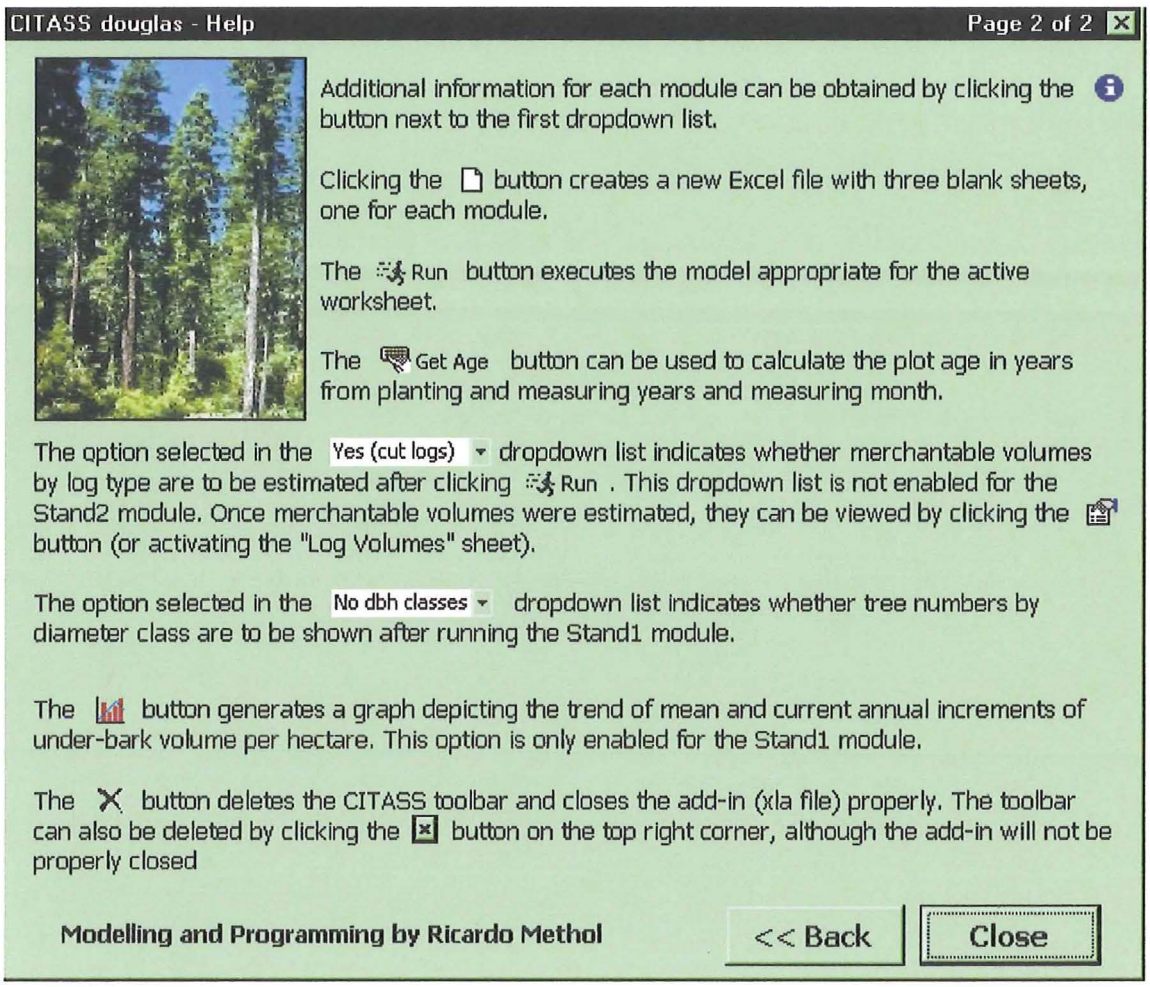
CITASS douglas is a Compatible Individual Tree and Stand Simulator for Douglas-fir plantations in the Central North Island of New Zealand. This simulator has three modules that can be accessed by (i) selecting the module name in the first dropdown list of the CITASS toolbar, or (ii) activating the worksheet corresponding to each module.

The "Stand1" module contains a diameter distribution model that allows thinning operations to be simulated. Merchantable volumes by log type at harvest and at each thinning can be estimated in this module. Up to 5 log types can be specified by setting log lengths and minumum under-bark small-end diameters (SEDmin). The algorithm used for simulate log cutting does not maximise recoverability or value. It simply tries to cut the log type inputted first (priority 1). If the diameter at the target tree position predicted by the taper function (tree taper table 136) is lower than the specified SEDmin, the algorithm tries to cut the log type inputted second (priority 2), and so forth.

The "Stand2" module contains a whole-stand model that can be used to project various plots at once, which may be useful for updating inventories. However, thinning simulations and merchantable volume estimations can only be performed in the "Stand1" module.

The "Tree" module contains an individual tree model fully compatible with the stand-level models in the other two modules. In this module, tree lists can be annually projected and various stand-level statistics (including merchantable volumes by log type) can be estimated.

Next >> Close



6) Specific information for each module can be obtained by clicking the "Info" button on the left-hand side of the CITASS toolbar.